3.6.2 Safety Procedures

The following cockpit rest period safety procedures were followed to minimize any interference with the safe operation of the aircraft.

1. Two crewmembers and two NASA observers were available while any one crewmember was resting.
2. The 20 min. recovery period was intended to allow sufficient time to return to full alertness and evaluate any concerns before re-entering the operational loop.
3. The potential for sleep inertia that might decrease performance was assessed (through inquiry by the NASA observers) before resuming flight duties.
4. A postrest update was provided on flight status and other relevant operational information before resuming flight duties.
5. The captain was to be alerted immediately upon first indication of any potential anomaly.
6. All rest periods were scheduled for completion at least 1 hr. before descent.
7. Safe, normal operation of aircraft was acknowledged as the highest priority, of course, and study procedures were not be permitted to interfere.

3.6.3 No-Rest/Control Group Procedures

Soon after TOC, the volunteer pilots in the NRG also identified a specific control period during the cruise portion of flight (see fig. 3: position a, b, or c). This served as a control period, and they followed the same procedures with a preparation time, 40 min. test period, and 20 min. "recovery" period when performance tests were administered. However, during the identified 40 min. control period, NRG pilots were instructed to continue their usual flight activities.

4.0 RESULTS

4.1 Subject Characteristics

Subject volunteer crews were randomly assigned to one of the two study groups. The NRG consisted of three crews totaling nine subjects. The RG consisted of four crews totaling 12 subjects. The mean age, mean years of experience, and sex of the volunteers are given in table 2. All of these factors were comparable between the two groups. One other field data collection trip, not included in this data set, was begun and then discontinued when rescheduling caused an alteration in the study trip schedule.

4.2 Pilot Choice of Rest Position

The procedures provided first choice of rest position to the landing pilot. Figure 4 shows the landing pilots' (for both captains and first officers [FOs]) choices for rest position a, b, or c and also the nonlanding pilots' choices. The main finding was that both captains and FOs generally chose the last rest position when they were landing the aircraft and rarely chose the first rest position. This result suggests that rather than rest early in the flight, when pilots may still be alert from layover sleep, the preferred strategy was to use the rest position later in the flight and closer to the landing.

Table 2. *Final study population subject characteristics*

	Sample size	Age (mean)	Experience (mean yr.)	Sex
No-rest group	**9**	**38.8**	**15.2**	**9M**
Captain	3	52.0	25.7	3M
FO	3	40.7	15.7	3M
SO	3	33.7	4.2	3M
Rest group	**12**	**38.7**	**13.4**	**11M 1F**
Captain	4	50.3	25.3	4M
FO	4	31.8	11.5	3M 1F
SO	4	34.0	3.5	4M
Total	**21**	**41.6**	**14.2**	**20M 1F**

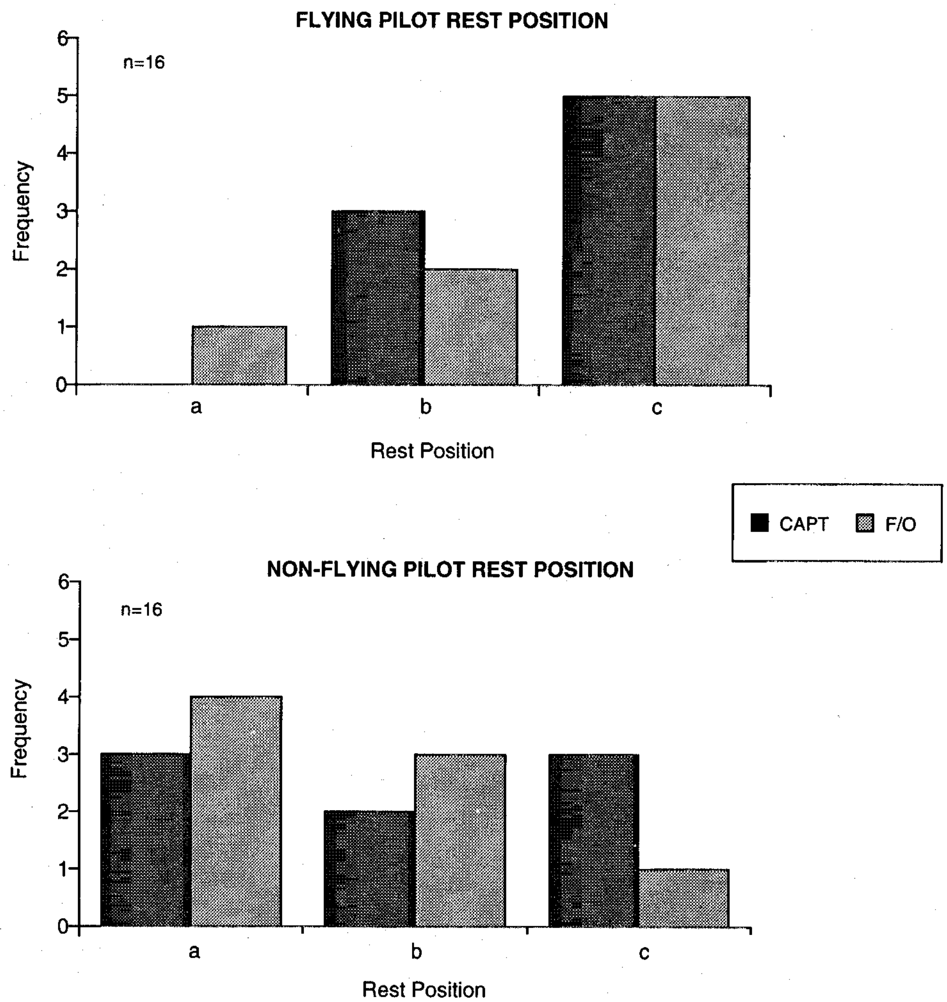


Figure 4. Choice of position for cockpit rest by landing vs. nonlanding pilot and captain vs. FO. (Landing pilot received first choice.)

4.3 Rest Period Sleep: EEG Findings

There were 12 subjects in the RG, each flying four segments, for a total of 48 rest periods. Physiological data were lost on three of these flight legs owing to a variety of factors. In view of the technical nature of the physiological recordings and the complexity of the operational environment in which the data were collected, a 6% data loss was considered minimal and acceptable.

First, the total number of sleep episodes that occurred is described. These data were then analyzed for six factors that describe the quantity and quality of sleep. These factors include the total amount of sleep (total sleep time), the sleep efficiency (total sleep time divided by the 40 min. rest period opportunity), the time to fall asleep (sleep latency), and the percentage of NREM sleep stages 1, 2, and slow-wave sleep. Each of these six factors was analyzed for overall descriptive summaries and also for (1) effects across study legs; (2) first-half study legs (study legs 1 and 2) vs. second-half study legs (study legs 3 and 4); (3) day (study legs 1 and 3) vs. night (study legs 2 and 4) leg differences; and (4) differences by flight position (captain vs. FO vs. SO).

The analyses for each of these six factors will be described in the text, with the significant findings highlighted graphically.

4.3.1 Total EEG Sleep Episodes

On 93% (42 out of 45) of the rest-period opportunities available for analysis, the RG subjects were able to sleep. There were three subjects who did not sleep on one flight leg each. One FO and one SO obtained no sleep on their fourth flight leg and one FO had no sleep on his third flight leg. All three of these subjects were able to sleep on the other three flight legs of their trips. A more detailed examination of these subjects will be presented later.

Two main analytical approaches were performed to examine the quantity and quality of sleep obtained during the planned rest period. First, the 42 rest periods in which sleep occurred were analyzed. The three no-sleep rest periods were not included in these analyses, for they would have artificially introduced an increased variability into the data set and potentially obscured meaningful results or suggested spurious findings.

In consideration of the potential sensitivity of these data, an even more conservative approach was used. The second analysis was conducted on the data from the six subjects, 24 rest periods, with complete physiological data. (Overall, three subjects had missing data due to equipment malfunctions and three subjects had one rest period with no sleep.) Statistical comparisons were conducted using analysis of variance. Those subjects for whom complete physiological data were obtained provided the most comprehensive representation of the physiological sleep that occurred across study flight legs. The importance of intersubject variability can be assessed in this data set.

Therefore, the RG (42) analysis represents the means and standard deviations for the 42 rest periods in which sleep occurred. The second analysis, RG (24), is the conservative ANOVA statistical comparisons based on the 24 rest periods that represent the six subjects with complete physiological data.

4.3.2 Total Sleep Time

The total sleep time was calculated as the total amount of sleep (in minutes) from sleep onset (defined as 1.5 min. of continuous sleep) until the final awakening. For the RG (42), the average total sleep obtained per rest period was 25.78 min. (SD = 9.58 min.). For the RG (24), the average total sleep obtained per rest period was 28.45 min. (SD = 6.28). Statistical analyses demonstrated that there were no significant differences related to trip legs (table 3), halves of the trip (table 4), day versus night flights (table 5), or by flight position (table 6). Overall, no significant findings emerged related to the average total sleep time.

Table 3. Average total sleep time by trip leg (min.)						
	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	26.76 (9.85)	28.24 (9.53)	19.12 (10.01)	28.89 (5.91)	2.44	.08
RG (24)	29.97 (7.46)	29.57 (6.21)	22.43 (4.59)	31.82 (2.20)	2.91	.07

Table 4. Average total sleep time by trip half (min.)				
	First trip half	Second trip half	F	p
RG (42)	27.70 (7.59)	21.73 (10.11)	3.30	.10
RG (24)	29.77 (4.16)	27.13 (2.65)	1.24	.32

Table 5. Average total sleep time by day vs. night (min.)				
	Day flights	Night flights	F	p
RG (42)	23.29 (7.31)	27.37 (7.83)	2.79	.12
RG (24)	26.20 (4.67)	30.69 (2.14)	3.18	.14

Table 6. Average total sleep time by flight position (min.)					
	Captains	First officers	Second officers	F	p
RG (42)	27.35 (1.68)	25.70 (5.75)	22.96 (9.75)	.45	.65
RG (24)	27.42 (2.05)	28.98 (2.14)	29.73 (1.84)	.86	.61

4.3.3 Sleep Efficiency: Total Sleep Time/40-Minute Rest Period

Sleep efficiency is the amount of time during an identified period that an individual is actually asleep. This parameter can reflect prior sleep loss when it results in more consolidated sleep and a higher sleep efficiency than might usually be expected. In the circumstances of this study, it was calculated by dividing the total sleep time by the 40 min. allowed for the rest period. Therefore, if a crewmember had slept the entire 40 min., the sleep efficiency would have been 100%. Obviously, this metric parallels the total sleep time results, and findings were not expected to vary from these. It provided some information, however, regarding the percentage of the rest period time spent asleep.

For the RG (42), the average sleep efficiency per rest period was 64.47% (SD = 23.94). For the RG (24), the average sleep efficiency per rest period was 71.12% (SD = 15.67). Statistical analyses demonstrated that there were no significant differences related to trip legs (table 7), halves of the trip (table 8), day versus night flights (table 9), or by flight position (table 10). Overall, no significant findings emerged related to the average sleep efficiency. As expected, this exactly parallels the total sleep time findings.

In a usual daytime nap, sleep efficiency would generally be in the 50%-55% range. Therefore, these results (64% and 71% sleep efficiency) may reflect accumulated sleep loss.

Table 7. Average total sleep efficiency by trip leg

	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	66.90%(24.59)	70.61%(23.84)	47.81% (25.00)	72.23% (14.76)	2.45	.08
RG (24)	74.95%(18.60)	73.90%(15.49)	56.10% (11.42)	79.52% (5.58)	2.91	.07

Table 8. Average total sleep efficiency by trip half

	First trip half	Second trip half	F	p
RG (42)	69.27%(18.96)	54.33% (25.27)	3.30	.10
RG (24)	74.43%(10.41)	67.81% (6.61)	1.24	.32

Table 9. Average total sleep efficiency by day vs. night

	Day flights	Night flights	F	p
RG (42)	58.22% (18.24)	68.45% (19.56)	2.81	.12
RG (24)	65.53% (11.63)	76.71% (5.31)	3.19	.13

Table 10. Average total sleep efficiency by flight position

	Captains	First officers	Second officers	F	p
RG (42)	68.39% (4.17)	64.29% (14.38)	57.40% (24.34)	.45	.65
RG (24)	68.54% (5.10)	72.48% (N=1)	74.30% (4.53)	.87	.50

4.3.4 Time to Fall Asleep: Sleep Latency

Sleep latency was defined as the time from the identified beginning of the 40 min. rest period to the first continuous 1.5 min. of sleep. For the RG (42), the average time to fall asleep per rest period was 5.55 min. (SD = 5.04 min.). For the RG (24), the average time to fall asleep per rest period was 4.10 min. (SD = 2.88 min.). Statistical analyses demonstrated that there were no significant differences related to trip legs (table 11), halves of the trip (table 12), day versus night flights (table 13), or by flight position (table 14). Overall, there were no significant findings related to the average sleep latency.

Table 11. Average sleep latency by trip leg (min.)						
	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	4.52 (3.29)	3.97 (2.73)	7.96 (6.65)	6.16 (6.61)	1.37	.27
RG (24)	3.15 (2.48)	3.82 (2.89)	5.12 (3.25)	4.33 (3.26)	0.50	.69

Table 12. Average sleep latency by trip half (min.)				
	First trip half	Second trip half	F	p
RG (42)	4.25 (2.98)	7.11 (6.51)	3.52	.07
RG (24)	3.48 (2.28)	4.73 (2.02)	1.20	.32

Table 13. Average sleep latency by day vs. night (min.)				
	Day flights	Night flights	F	p
RG (42)	6.02 (4.41)	4.92 (3.49)	.62	.45
RG (24)	4.13 (2.01)	4.08 (2.71)	.002	.97

Table 14. Average sleep latency by flight position (min.)					
	Captains	First officers	Second officers	F	p
RG (42)	5.83 (3.11)	4.42 (3.67)	7.10 (4.72)	0.48	.64
RG (24)	4.34 (1.11)	1.25 (N=1)	5.18 (0.71)	5.37	.10

4.3.5 Percent NREM Stage 1 Sleep

NREM stage 1 sleep is the lightest sleep stage. This metric portrayed the percentage of total sleep time spent in NREM stage 1 sleep and provided some indication of the depth of the sleep obtained. For the RG (42), the average NREM stage 1 percent per rest period was 30.28% (SD = 22.50). For the RG (24), the average NREM stage 1 percent per rest period was 24.75% (SD = 15.52). There was a significant effect related to trip legs (table 15), but there were no significant findings related to halves of the trip (table 16) or flight position (table 17).

Post hoc analyses of the RG(24) were performed to understand more fully the significant contribution by trip leg. Two significant post hoc comparisons emerged. The average NREM stage 1 sleep percent on leg 1 (23.10%) was significantly greater than the leg 4 (10.00%) average NREM stage 1 sleep percent ($F_{1,5} = 13.58$, $p = .01$) (A p value equal to .01 indicates that there is a 99% confidence that this is a significant finding due to trip leg and would only occur by chance 1 time in a 100). Also, the average NREM stage 1 sleep percent on leg 3 (37.00%) was significantly greater than the leg 4 (10.00%) average NREM stage 1 sleep percent ($F_{1,5} = 36.76$, $p = .002$).

The average NREM stage 1 sleep percent on the day legs (1 and 3) was significantly greater than the NREM stage 1 sleep percent on the last night leg (leg 4).

There was also a significant effect for average NREM stage 1 sleep percent related to day versus night flights (table 18). The average NREM stage 1 percent for day flights (legs 1 and 3) was greater than the average NREM stage 1 percent for night flights (legs 2 and 4). There was a significant effect for the RG (42) subjects and a similar statistical trend in the more conservative analysis for the RG (24) subjects.

Table 15. Average NREM stage 1 sleep percent by trip leg

	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	28.27%(15.61)	28.22%(20.79)	47.63% (29.58)	16.21%(11.19)	3.90	.02*
RG (24)	23.10%(11.37)	28.90%(19.17)	37.00% (10.99)	10.00% (5.04)	4.63	.02*

* p < .05.

Table 16. Average NREM stage 1 sleep percent by trip half

	First trip half	Second trip half	F	p
RG (42)	28.23% (9.80)	36.84% (30.29)	1.01	.34
RG (24)	26.00% (7.80)	23.50% (6.59)	0.56	.49

Table 17. Average NREM stage 1 sleep percent by flight position

	Captains	First officers	Second officers	F	p
RG (42)	24.25% (6.11)	28.50% (8.69)	40.85% (18.29)	1.99	.19
RG (24)	26.53% (4.98)	15.95% (N=1)	26.48% (5.80)	1.68	.32

Table 18. Average NREM stage 1 sleep percent by day vs. night

	Day flights	Night flights	F	p
RG (42)	37.30% (13.47)	25.32% (16.77)	8.00	.02*
RG (24)	30.05% (5.65)	19.45% (9.71)	6.04	.06

* p < .05.

4.3.6 Percent NREM Stage 2 Sleep

NREM stage 2 sleep is a deeper sleep stage than NREM stage 1. It is the predominant sleep stage during nocturnal sleep, comprising about 50% of total sleep time. This metric portrays the percentage of total sleep time spent in NREM stage 2 sleep. For the RG (42), the average NREM stage 2 percent per rest period was 61.65% (SD = 21.63). For the RG (24), the average NREM stage 2 percent per rest period was 67.30% (SD = 17.66). Statistical analyses demonstrated that there were no significant differences related to trip legs (table 19), halves of the trip (table 20), day vs. night flights (table 21), or by flight position (table 22). Overall, no significant findings emerged related to the average NREM stage 2 percent.

Table 19. Average NREM stage 2 sleep percent by trip leg

	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	64.91% (13.94)	62.85% (19.22)	52.15% (29.32)	66.38% (23.37)	0.88	.46
RG (24)	66.30% (9.98)	70.22% (18.57)	63.00% (10.99)	69.67% (28.80)	0.38	.77

Table 20. Average NREM stage 2 sleep percent by trip half

	First trip half	Second trip half	F	p
RG (42)	64.27% (13.18)	55.00% (28.83)	1.13	.31
RG (24)	68.26% (13.71)	66.33% (17.19)	0.23	.65

Table 21. Average NREM stage 2 sleep percent by day vs. night

	Day flights	Night flights	F	p
RG (42)	58.44% (11.32)	63.13% (19.89)	0.99	.34
RG (24)	64.65% (7.29)	69.94% (23.27)	0.52	.50

Table 22. Average NREM stage 2 sleep percent by flight position

	Captains	First officers	Second officers	F	p
RG (42)	61.68% (12.70)	66.04% (12.59)	54.40% (16.97)	0.68	.53
RG (24)	64.03% (14.45)	83.62% (N=1)	64.04% (18.69)	0.63	.59

4.3.7 Percent NREM Slow-Wave Sleep

NREM slow-wave sleep is the deepest sleep. It is a combination of both NREM stages 3 and 4 and reflects the number of EEG delta waves. This metric portrays the percentage of total sleep time spent in NREM slow-wave sleep and provides some indication of the depth of the sleep obtained. For the RG (42), the average NREM slow-wave sleep percent per rest period was 8.07% (SD = 16.22). For the RG (24), the average NREM slow-wave sleep percent per rest period was 7.96% (SD = 18.01). There were no significant differences related to trip legs (table 23), halves of the trip (table 24), or flight position (table 25).

There was a significant effect for the average NREM slow-wave sleep percent for day versus night flights (table 26). The average NREM slow-wave sleep percent for day flights (legs 1 and 3) (4.3%) was less compared to the average NREM slow-wave sleep percent for night flights (legs 2 and 4) (11.6%).

Table 23. Average NREM slow-wave sleep percent by trip leg

	Leg 1	Leg 2	Leg 3	Leg 4	F	p
RG (42)	6.83% (12.10)	8.94% (16.48)	0.22% (0.70)	17.40% (25.07)	1.93	.14
RG (24)	10.62% (15.23)	0.88% (2.16)	0.00% (0.00)	20.33% (30.46)	2.47	.10

Table 24. Average NREM slow-wave sleep percent by trip half

	First trip half	Second trip half	F	p
RG (42)	7.51% (10.62)	8.15% (12.13)	0.05	.84
RG (24)	5.75% (7.31)	10.17% (15.23)	1.65	.26

Table 25. Average NREM slow-wave sleep percent by day vs. night

	Day flights	Night flights	F	p
RG (42)	4.27% (7.24)	11.56% (14.05)	7.57	.02*
RG (24)	5.31% (7.61)	10.61% (15.02)	2.51	.17

* $p < .05$.

Table 26. Average NREM slow-wave sleep percent by flight position

	Captains	First officers	Second officers	F	p
RG (42)	14.08% (14.69)	5.46% (6.40)	4.74% (9.23)	0.95	.42
RG (24)	9.46% (13.99)	0.43% (N=1)	9.48% (12.87)	0.18	.84

Thus, there was one significant finding that emerged for these factors. As a group, crewmembers had a higher percentage of NREM slow-wave sleep during night flights than on day flights. This suggests that deeper sleep occurred on night flights than on day flights.

Figure 5 presents the NREM stage 1, stage 2, and slow-wave sleep percentages of total sleep time for day versus night flights. This portrays the day flights with more light sleep and less deep sleep and the night flights with less light sleep and more deep sleep.

No REM sleep was observed in any of the rest period sleep episodes.

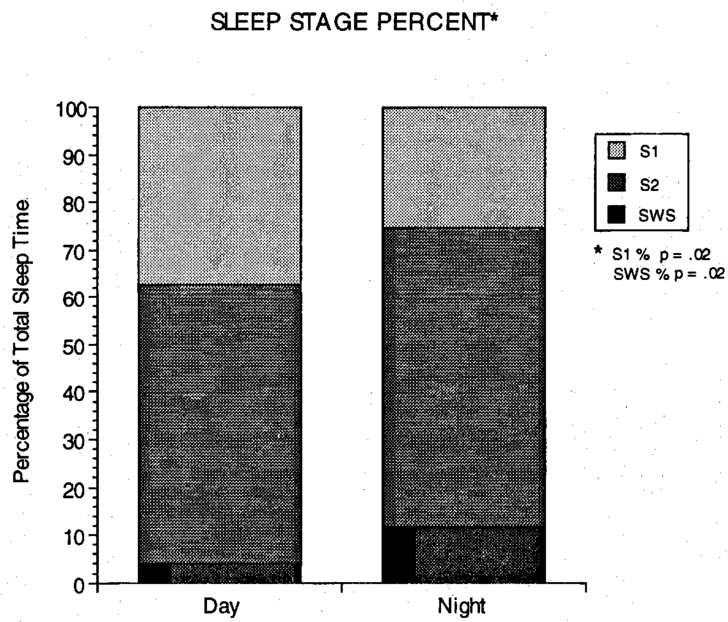


Figure 5. Percentage of total sleep time in stage 1, stage 2, and slow-wave sleep, by day vs. night flights.

4.3.8 RG Subjects With No Sleep

As indicated previously, there were three pilots who did not sleep on three separate rest period opportunities. Figure 6 portrays the sleep obtained on the other three legs of their trip schedules for these subjects. Each column indicates the total amount of sleep, composed of the total stage 1 sleep (TS1), total stage 2 sleep (TS2), and total slow-wave sleep (TSWS). Several points can be noted from these data. First, all three of the no-sleep episodes occurred later in a trip, with two of three on the fourth leg. Examination of the figure suggests that one FO and one SO generally slept below the RG average amounts. In particular, the SO demonstrated a relatively poor ability to obtain sleep on all but the first trip leg. These patterns and the subjects' inability to sleep on these three occasions highlight the complexity of this situation. There are a variety of factors that may have played a role in their inability to sleep, for example, individual differences, personality characteristics, circadian factors, or different sleep patterns. It is important to note that these individuals were able to obtain sleep on all other flight legs. A more detailed examination of factors that may have led to these no-sleep episodes is planned.

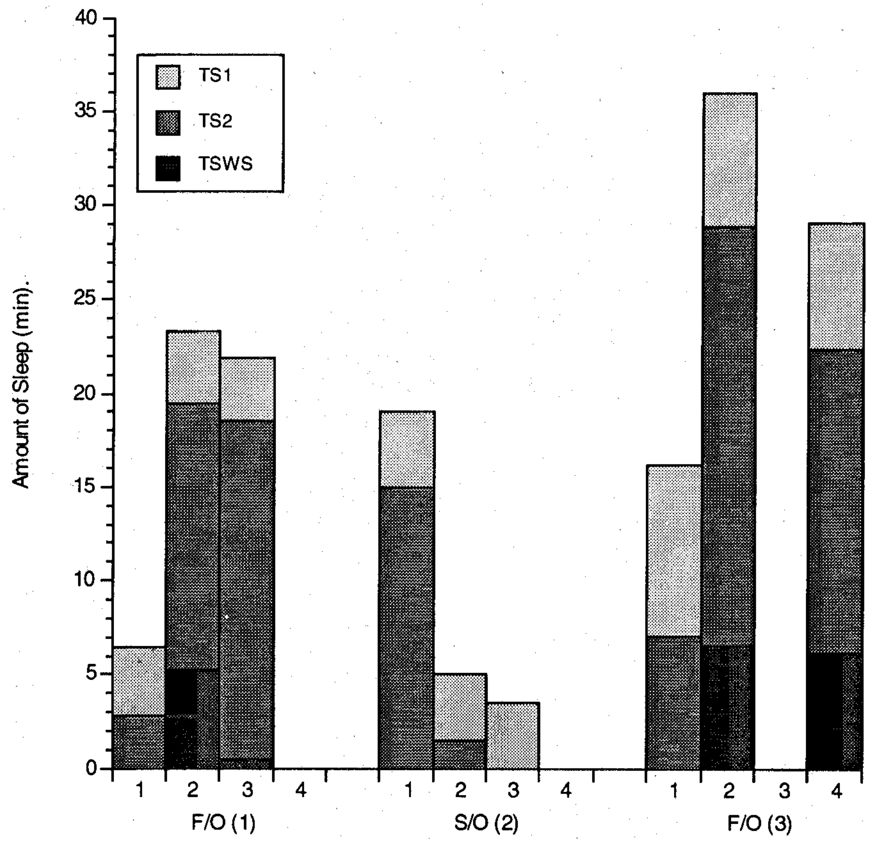


Figure 6. Percentage of total sleep time in stage 1, stage 2, and slow-wave sleep, for each of the three subjects in RG who did not sleep on one of the four trip legs; data given for the three legs that did include sleep.

4.3.9 NRG Subjects With Sleep During the Control Period

An interesting finding emerged from examination of the control period in the NRG subjects. This group underwent the exact same measurement and performance evaluation procedures as the RG; however, during the pre-identified control period, the NRG subjects were instructed to conduct their usual flight activities. Analysis of the EEG recordings for the 40 min. control periods for the NRG demonstrated that four NRG subjects fell asleep on a total of five occasions (one subject fell asleep during two different control periods). Four of the nine NRG subjects (44%) had at least one episode of spontaneous sleep during the control period. The total sleep (in minutes) for the five episodes is shown in figure 7. Although there were a couple of brief sleep episodes, two of the periods were over 10 min. long. Only NREM stage 1 and NREM stage 2 sleep occurred during these episodes; there was no deep NREM slow-wave or REM sleep.

4.4 Psychomotor Vigilance Task Performance

The 21 crewmembers who participated in the study each performed between 180 and 190 min. of the psychomotor vigilance task (PVT), for a total of 63 hr. of performance assessment (over 26,000 reaction times). For all four flight legs of the study, a 10 min. PVT trial was administered 1-2 hr. before each flight (preflight trial) and three times during the cruise portion of each flight (in-flight trial 1 was before the rest or control period, in-flight trial 2 was immediately after the rest or control period, in-flight trial 3 was before TOD). The PVT was also administered 1-2 hr.

following each flight (postflight), with the exception of study flight leg 4 (NRT to LAX) due to logistical problems. For this reason, analysis of PVT data was conducted in two basic ways. First, a 2 by 4 by 4 (rest/control conditions x study flight legs 1, 2, 3, 4 x preflight trial and in-flight trials 1, 2, 3) mixed-model analysis of variance (ANOVA) was conducted on each performance parameter. Second, a separate two-way mixed model ANOVA was carried out within each flight leg, utilizing the postflight trial on all but the fourth leg.

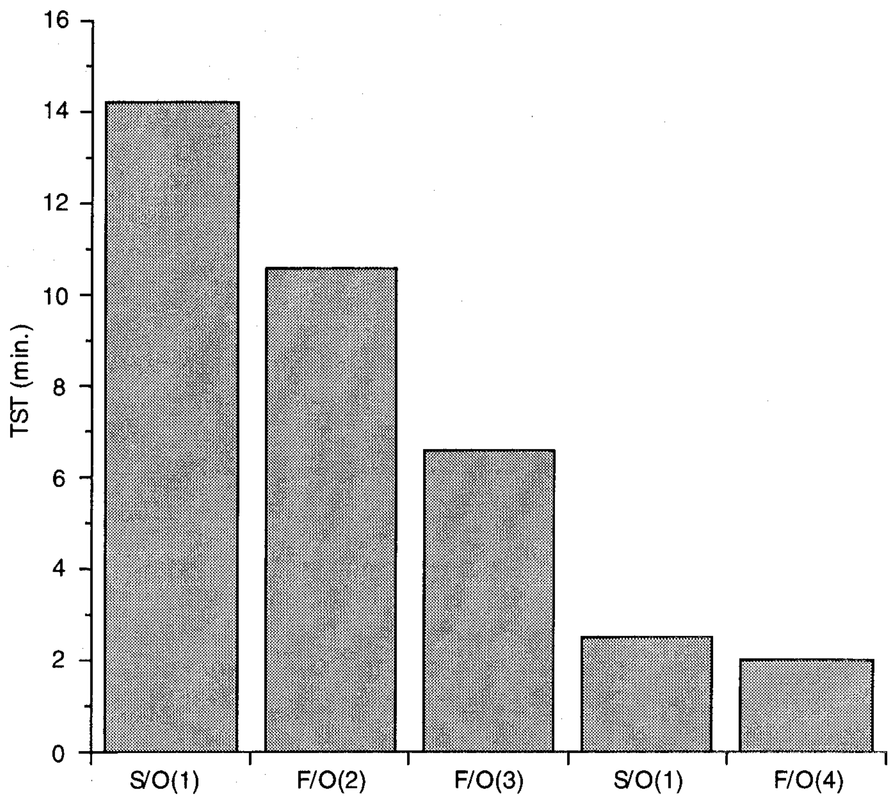


Figure 7. Total sleep time for the four NRG control subjects who fell asleep (for a total of five sleep episodes) during the 40 min. test period.

As described previously, the PVT data were analyzed for response slowing (median reaction time), lapse frequency, lapse duration, optimum response time, and vigilance decrement. Rather than review all of the PVT data, results for the median reaction time (response slowing), lapse frequency, and vigilance decrement are presented here. The results from the lapse duration and optimum response time provide similar findings and are presented in the appendix.

4.4.1 PVT Response Slowing (Median Reaction Time)

A characteristic feature of fatigue is the slowing of response output on cognitive tasks (ref. 6). Response slowing across PVT trials was assessed by determining the median reaction time (RT) per trial, to prevent a disproportionate influence from long-duration lapses. (For a discussion of increased performance variability caused by sleep loss, and the statistical approach to handle this variability, see ref. 58). Figure 8 shows the average of median RTs for the no-rest and rest groups for each trial of each study flight leg. The NRG displays far greater range of average responses across flight legs and trials than the RG, with response slowing especially evident on the third in-flight performance trial on study flight legs 2 and 4. The three-way ANOVA confirmed this observation. There were significant main effects for condition ($F_{1,19} = 9.19$, $p < .007$), flight leg

($F_{3,57} = 5.18$, $p < .003$), and trial ($F_{3,57} = 12.93$, $p < .0005$). There were significant interactions for condition by flight leg ($F_{3,57} = 3.38$, $p < .025$), and condition by trial ($F_{3,57} = 5.17$, $p < .003$), as well as for flight leg by trial ($F_{9,171} = 4.90$, $p < .0005$). The F-ratio for the three-way interaction was not significant ($F_{9,171} = 0.87$).

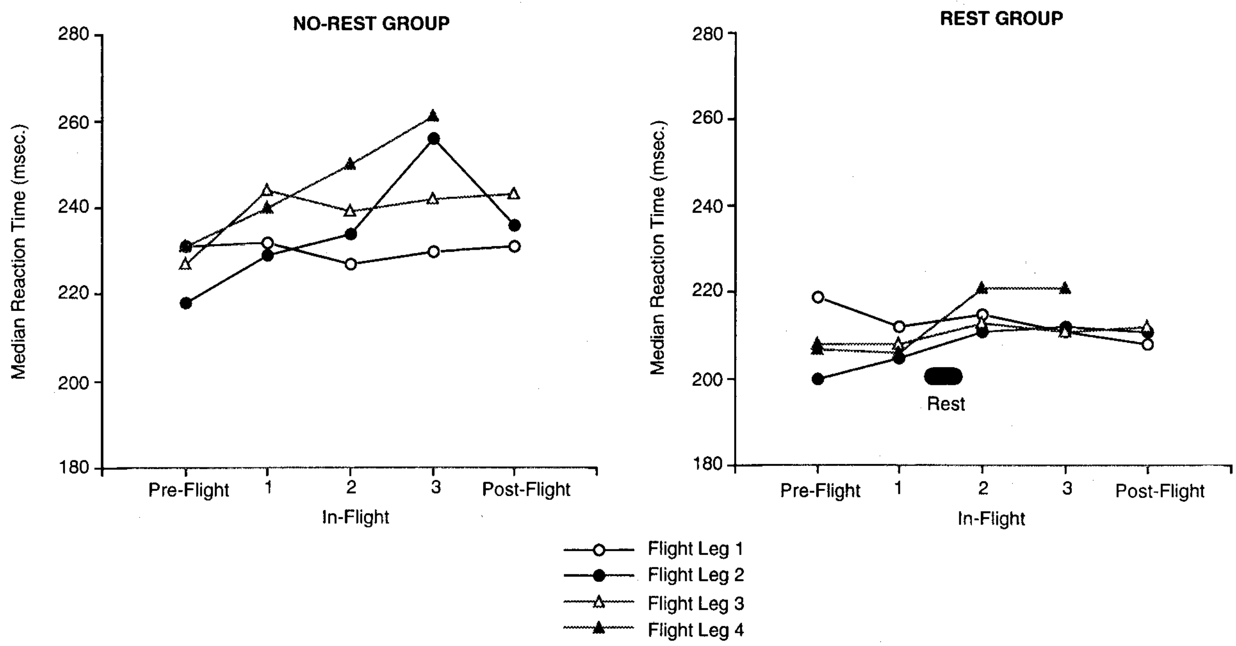


Figure 8. Median RT for each 10 min. PVT trial for RG and NRG across each flight leg; data points are averages of the medians within each group, and increases indicate poorer performance.

Two-way ANOVAs further clarified these effects. There were no significant main effects or interactions on study flight leg 1—the two groups were performing comparably at this time. However, on study flight legs, 2, 3, and 4, the NRG exhibited significantly more response slowing than the RG (main effect for condition: leg 2 $F_{1,19} = 11.73$, $p < .003$; leg 3 $F_{1,19} = 12.65$, $p < .002$; leg 4 $F_{1,19} = 8.92$, $p < .008$). Figure 9 illustrates this effect using data from the first and last study flight legs (1 and 4, respectively). The NRG displays a steady increase in median RT across flight leg 4 relative to flight leg 1, with differences becoming statistically significant midway and late in flight. Such changes are not evident in the RG.

Figure 10 displays the difference at each trial time-point between the two groups for data averaged across the four study flight legs. The preflight difference is not statistically significant, but on average the NRG was 10%-16% slower than the RG during the in-flight trials and during the postflight trial. The maximum difference occurs for in-flight trial 3 prior to TOD.

4.4.2 PVT Lapse Frequency

The most widely known effect of sleep loss on performance is lapsing, which refers to a period of response delay (block or gap), resulting in progressive unevenness (increased variability) in the performance of a fatigued subject. (For a complete discussion of this phenomenon and the lapse hypothesis, see refs. 6, 58.) Lapses have been shown to be associated with microsleep events in the EEG (refs. 6, 64-67). In the last two decades, there have been many studies showing that sleep-based fatigue results in lapsing and increased performance variability on short-duration RT tasks involving sustained attention (refs. 13, 47, 48, 51-55). Although a number of definitions have been used, lapses have most often been defined as RTs twice as long as the baseline RT average (ref. 68). For a simple visual PVT of the kind used in this study, this value is 500 msec. (2 x 250 msec.). Thus, a lapse was defined as any RT longer than a half a second.

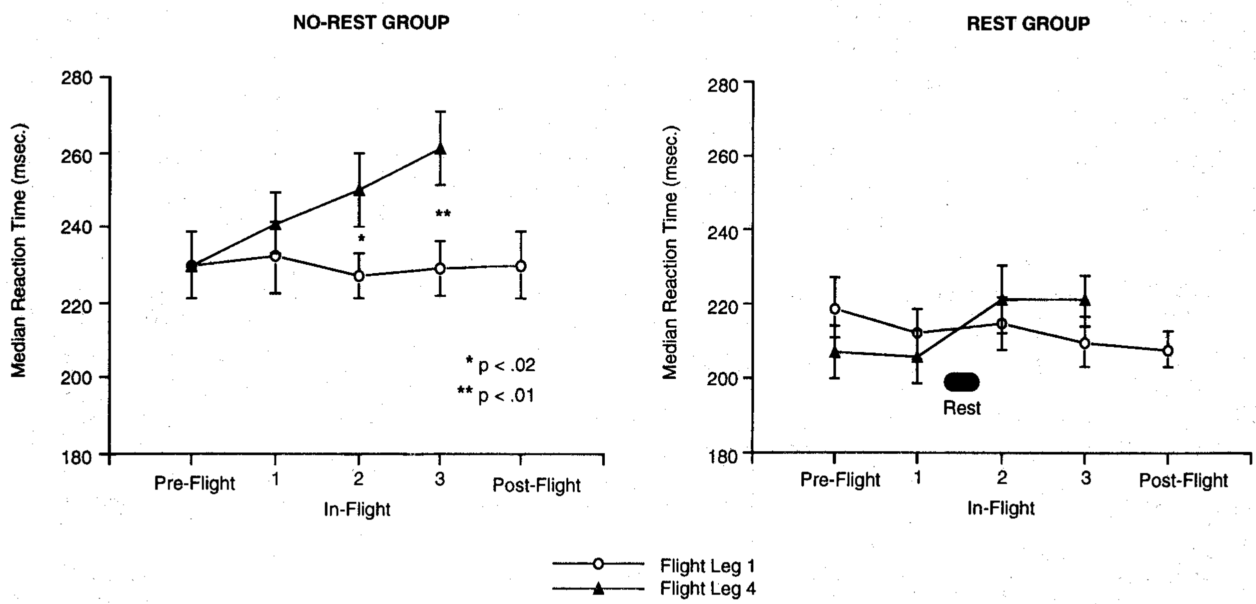


Figure 9. Median RT for each 10 min. PVT trial for both RG and NRG for day-flight leg 1 and night-flight legs 4; data points are averages (standard error bars) of the medians within each group. Increases indicate poorer performance; asterisks indicate significant differences with group by paired t-tests at specific time points.

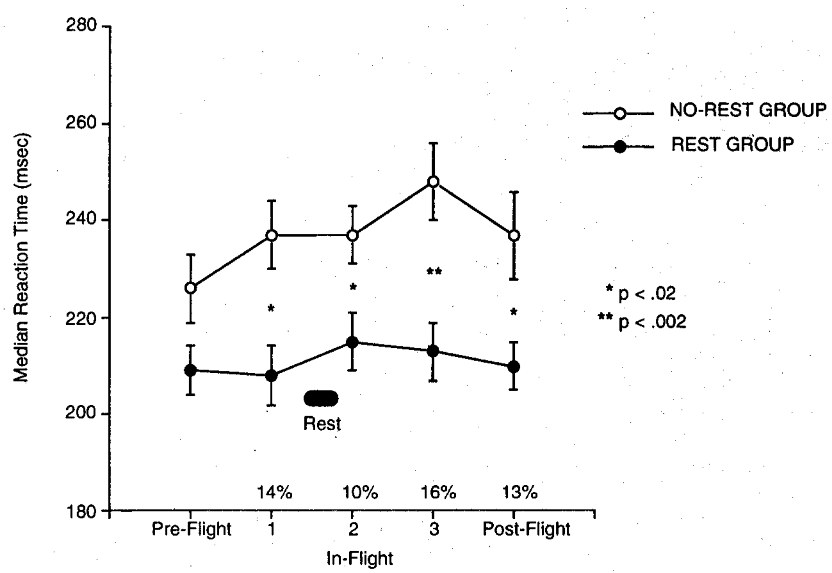


Figure 10. Median RT for each 10 min. PVT trial for both RG and NRG collapsed across all four flight legs; data points are averages (S.E. bars) of the medians within each group. Percentages indicate differences between group means at PVT trial times. Increases indicate poorer performance; asterisks indicate significant differences between groups by independent t-tests at specific time points.

There were a total of 283 lapses recorded for all 21 crewmembers in the study, representing about 1% of all PVT responses. As expected, lapses rarely occurred on PVT trials early in the study, when crews had fewer circadian disruptions and had accumulated less sleep debt. For all 21 crewmembers combined, the total number of lapses for the three in-flight PVT trials on day-flight leg 1 (HNL-OSA) was 20, which is only 7% of all lapses, and 10% of all in-flight lapses observed in the study. Lapses increased in frequency as crews progressed through the study, but the effect was more pronounced in the NRG (58% of all lapses) and on night-flight legs for both groups (60% of all lapses). Figure 11 shows the total number of lapses that occurred during PVT trials completed in the cruise portion of all four flight legs combined. The NRG had more total lapses in flight (N = 124) than the RG (N = 81), even though there were three more crewmembers in the RG than in the NRG. Moreover, the increase in lapses in the NRG is especially evident during in-flight performance trials 2 and 3, suggesting that the RG nap after trial 1 reduced the likelihood of increased lapsing later in the flight.

There were, however, broad individual differences in lapse frequency within each group. Five of nine NRG crewmembers had 10 or more in-flight lapses. Two of these crewmembers (an FO and a captain on different flights) had a disproportionately high total number of in-flight lapses (45 and 33, respectively), which together accounted for 38% of all in-flight lapses in the NRG. In contrast, only three of the 12 RG crewmembers totaled 10 or more lapses in-flight, and none had more than 14 lapses. Remarkably, five RG crewmembers (three captains and two SOs), as well as three NRG crewmembers (2 FOs and 1 SO), accumulated no more than four in-flight lapses on the PVT during the entire study, which is a rate of less than or equal to one lapse per flight.

Lapses, as more serious performance failures, require some consideration before statistical analysis because they comprise only a very small portion of all PVT responses and because there were such large differences in lapse frequency between individual crewmembers. Therefore, before conducting the ANOVAs, a square root transformation was used on the frequency count of the number of lapses to remove the proportionality between the mean and the variance (ref. 69). The results are presented in Figures 12-14. This analysis refers only to the number of lapses, without regard for their duration.

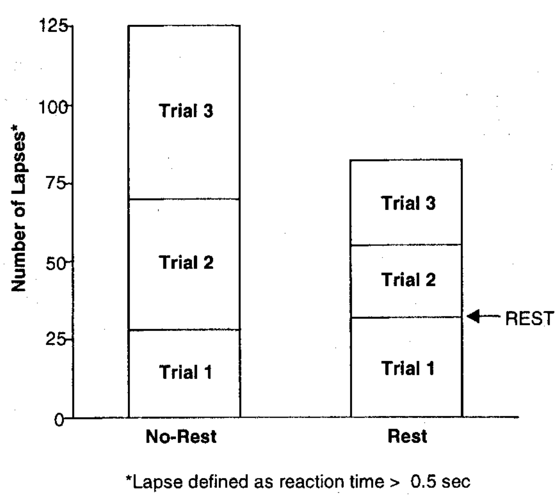


Figure 11. Cumulative number of raw lapses (RTs > 500 msec.) for PVT trials completed during the cruise portion of all four flight legs for the RG and NRG. The cockpit nap (rest) occurred between in-flight PVT trials 2 and 3. Increases indicate poorer performance.

Figure 12 displays the average number of transformed lapses for the NRG and RG for each trial of each study flight leg. The NRG averaged increasing numbers of lapses across flight legs and trials relative to the RG, particularly during the third in-flight performance trial (near TOD) on study night-flight legs (2 and 4). There was, however, considerable difference between the two groups in variability, not fully obviated by the transformation. This was reflected in the three-way ANOVA. There were significant main effects for flight leg ($F_{3,57} = 4.81$, $p < .005$) and trial ($F_{3,57} = 4.14$, $p < .01$), but not for condition. There was no interaction for condition by flight leg, and only a trend for a condition by trial interaction ($F_{3,57} = 2.18$, $p < .10$). The flight leg by trial interaction was significant ($F_{9,171} = 2.77$, $p < .005$). The three-way interaction was not.

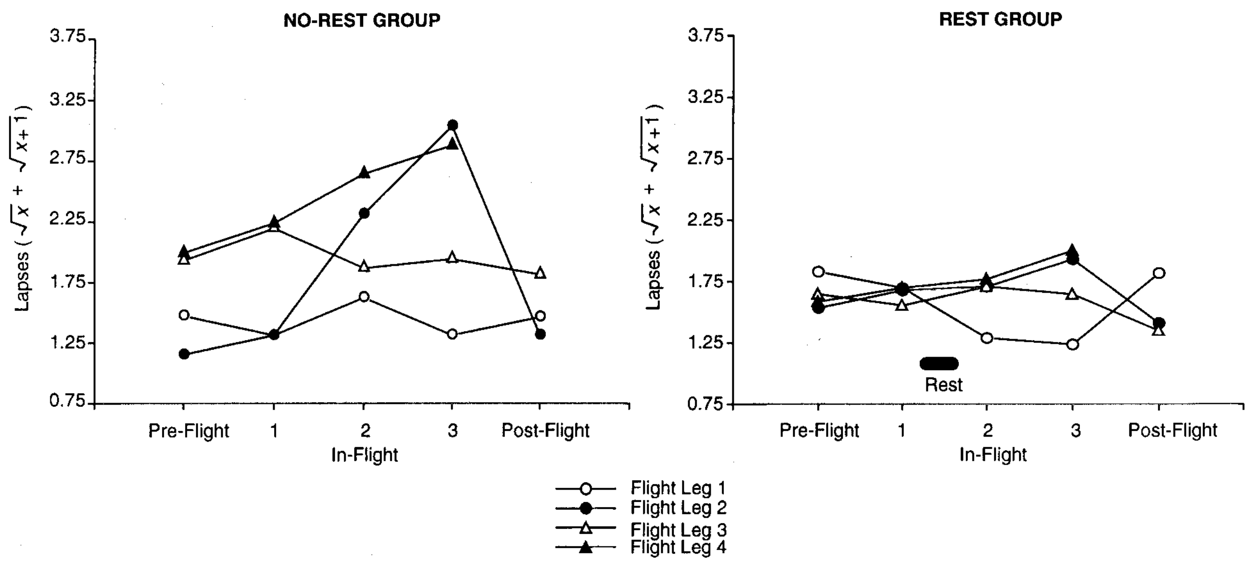


Figure 12. Mean number of transformed lapses (RT > 500 msec.) for each 10 min. PVT trial for both RG and NRG across each flight leg. Increases indicate poorer performance.

Two-way ANOVAs performed for data within each flight leg revealed that there were no main effects or interactions for day-flight legs (1 and 3) and for night-flight leg 4 (recall that this final leg did not have a postflight trial, reducing the degrees of freedom available). Night-flight leg 2 was associated with a significant condition by trial interaction ($F_{4,76} = 2.54$, $p < .05$). The NRG averaged increasing numbers of lapses during the flight relative to the RG. However, as noted above, not everyone in the NRG displayed increased lapsing on night flights, which accounts for the far greater variance around this group's mean on night-flight leg 4 (see fig. 19). As shown in figure 13, both groups had more lapses at TOD on night-flight leg 4 than at TOD on night-flight leg 1, but the increase from flight leg 1 to 4 is twice as large in the NRG as it is in the RG.

Figure 14 displays the difference at each trial time-point between the two groups for lapse frequency data averaged across the four study flight legs. At preflight and early in-flight (still pre-rest period), there were no differences between the RG and NRG in the average number of lapses or in the intersubject variability of lapsing. After the rest period, however, there were 30% more lapses during the two in-flight performance trials, and, more important, there was significantly greater variability ($p < .002$) at each time-point among NRG crewmembers (i.e., the intragroup variability of the NRG exceeded the intergroup variability). Thus, there was no sharp rise in lapses later in the flight for the RG. However, in the NRG, some subjects showed no increase in lapsing, whereas others had dramatic increases. It can be concluded that one benefit of the nap was to prevent some RG crewmembers from lapsing, especially during night flights.

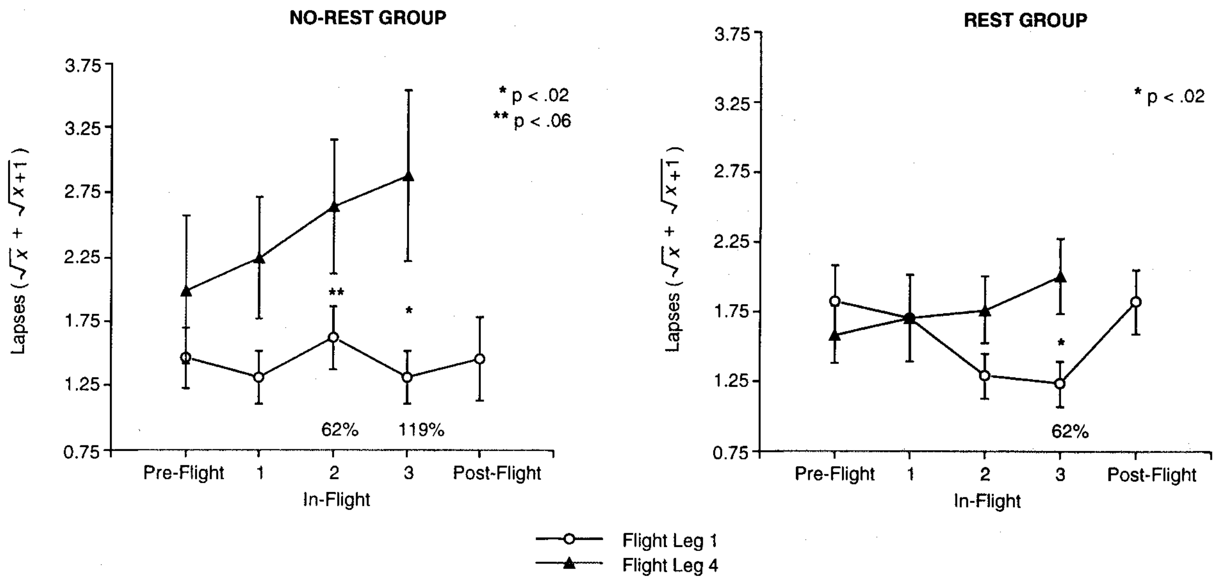


Figure 13. Mean (S.E.) number of transformed lapses (RT > 500 msec.) for each 10 min. PVT trial for both RG and NRG for day-flight leg 1 and night-flight leg 4. Increases indicate poorer performance. Percentages indicate differences between means of flight legs within each group at PVT trail times. Asterisks indicate significant differences with group by paired t-tests at specific time points.

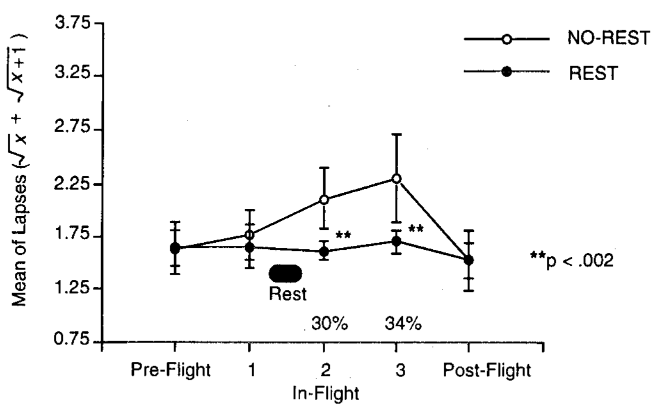


Figure 14. Mean (S.E.) number of transformed lapses (RT > 500 msec.) for each 10 min. PVT trial for both RG and NRG collapsed across all four flight legs. Increases indicate poorer performance. Percentages indicate differences between groups' means at PVT trial times. The asterisks indicate differences between groups in variance at in-flight trials 2 and 3. Asterisks indicate significant differences with group by paired t-tests at specific time points.

4.4.3 PVT Vigilance Decrement

The rate at which a response declines as a function of being repeated, or of time-on-task, reflects vigilance decrement. This same concept has been used in various literatures to define fatigue and habituation. There is a rich tradition of experimentally assessing changes in performance with time-on-task, and much of the classic literature on sleep deprivation effects used this approach (for reviews see refs. 6, 58). There is strong experimental evidence that sleep-based fatigue results in accelerated decrements in responding across the 10 min. PVT (ref. 50). This observation has proved to be theoretically valuable in understanding the role of environment in fatigue-based deficits (ref. 19). In fact, Dinges has suggested that his time-on-task PVT performance metric, which is the vigilance decrement function, is best conceptualized as an index of "fatigueability." This approach (ref. 50) has also provided a common metric by which to compare the magnitude of fatigue-based performance impairments between laboratory and field research, as seen in figure 15.

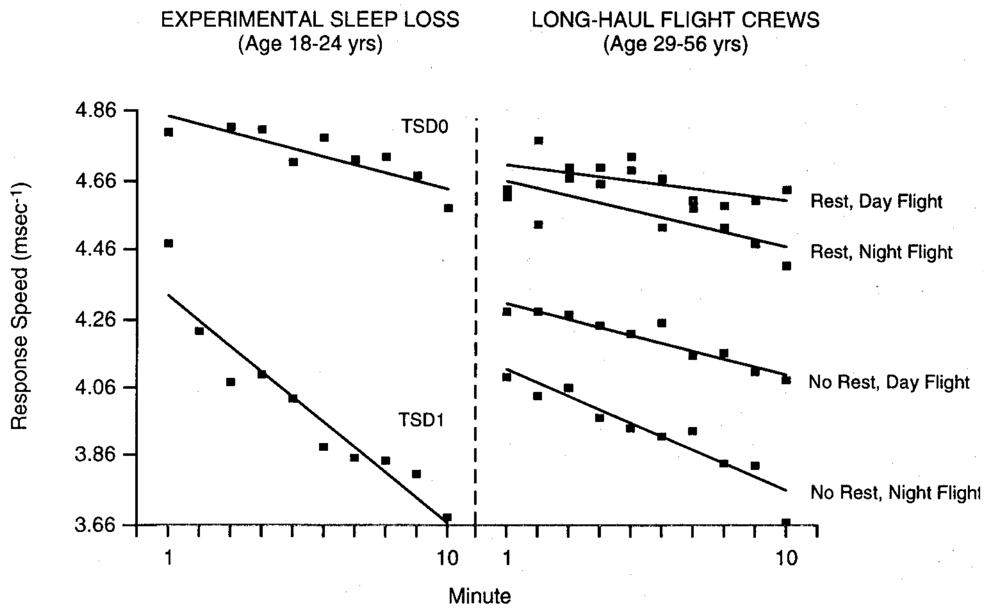


Figure 15. Vigilance decrements during PVT performance trial. Linear regression lines fitted by the method of least squares to the min.-by-min. average response speed across the 10 min. PVT. Data in right hand panel are from in-flight performance trials for RG and NRG during day-flight legs (mean of legs 1 and 3) and night-flight legs (mean of legs 2 and 4). For crews in the current study, only RTs from the second (mid-flight) and third (near TOD) in-flight PVT trials were used for each of the four lines. Data in left-hand panel are for comparison purposes from a study of nine healthy young adults performing the PVT during a day following a normal night of sleep (TSD0) and following 1 night without sleep (TSD1). Each regression line was fitted to the average performance across 10 min. Decreases indicate poorer performance.

The right half of figure 15 shows the linear regression lines fitted to the minute-by-minute average response speed across the 10 min. PVT for the RG and NRG on day-flight legs (mean of legs 1 and 3) and night-flight legs (mean of legs 2 and 4). (Note: Because of a 1/RT statistical transformation, a downward deflection indicates poorer performance.) Only RTs from second (mid-flight) and third (near TOD) in-flight PVT trials were used for each of the four lines, and data were averaged within subjects and then between subjects for comparable time-points to generate these functions (hence each regression line in fig. 15 represents the function fitted to the minute-by-minute averages, not the average of the functions for each crewmember). Linear regression lines were fitted to the transformed data by the method of least squares.

On the left-hand side of figure 15 are data from college students performing the PVT during a day following a normal night of sleep (TSD0 = total sleep deprivation 0, i.e., 3-17 hr. awake), and following 1 night without sleep (TSD1 = total sleep deprivation 1 night; i.e., 18-42 hr. awake). The mean vertical difference between lines (or their y-intercepts) reflects the overall response slowing engendered by fatigue from sleep loss and night flights. The slopes of the regression equations provide an estimate of the fatigueability of crewmembers. In all cases in figure 15, the correlations of fit are statistically significant ($p < .05$ or higher), and range between .67 (rest night flight) and .95 (no-rest day flight).

As evidenced in earlier figures, the RG subjects had a higher mean response speed than the NRG subjects, and despite a considerable difference in age, their in-flight mean performance level (y-intercept) and fatigueability (slope) was near to that of healthy young adults who had not been sleep deprived. There is a tendency, evident in figure 15, for the RG subjects to be slightly slower and more fatigueable on night flights than on day flights. The difference is trivial, however, compared with how much better their average performance was relative to the NRG subjects, and compared to the average difference between day and night flights within the NRG. The NRG fatigueability function fitted to average data (slope = -.039) is less steep than that of the average laboratory subjects deprived of a night's sleep (TSD1 slope = -.073). However, the combined lower y-intercept and steeper slope suggest that during night flight, the NRG crewmembers were approaching a fatigue level that could be characterized as undesirable.

The fatigueability functions in figure 15 are based on regressions fitted to average data. Therefore, they do not indicate intersubject variability, or whether the greater slope for the NRG on night flights is statistically different from their day flights. Also, they do not determine how these differences compare with the day and night slopes for the RG. To obtain these answers, regression lines were fitted to the transformed minute-by-minute data for each individual crewmember. Those crewmembers in either group who had a y-intercept difference between day flight and night flight of more than 0.2 (which is between 8 and 18 msec. in raw RT) were excluded from the analyses. The reason for this criterion was to assess differences in fatigueability (slope), given roughly comparable initial levels of functioning (y-intercept). Application of this criterion reduced the NRG from nine to seven subjects, and the RG from 12 to 8 subjects. Despite the loss of degrees of freedom, this approach yielded an important observation.

Although there continued to be significant mean differences in y-intercepts between the NRG and RG, there were no significant differences within either subgroup in y-intercepts for day and night flights (which was the purpose of applying the criterion). The average (SD) regression slope for the eight RG crewmembers during day flight was -.026 (.025); during night flight it was -.023 (.023). The average regression slope for the seven NRG crewmembers during day flight was -.022 (.012); during night flight it was -.047 (.018). A two-way repeated-measures ANOVA yielded a significant interaction ($F_{1,13} = 6.94$, $p < .021$), but no main effects. The night-flight slope for the NRG was significantly steeper than its day-flight slope ($t = 4.29$, $p < .002$), and steeper than the RG night slope ($t = 2.18$, $p < .048$).

Thus, given comparable initial levels of performance, only the NRG crewmembers displayed greater fatigueability on night flights than on day flights. This suggests that one outcome of the cockpit nap was to prevent increased fatigueability on the night flight. The magnitude of the difference is remarkable. During night flight, the average NRG response speed declined with time-on-task (mean slope = -.047) twice as fast as that of the RG (mean slope = -.023). This result is more noteworthy when one considers that the two NRG crewmembers excluded from the analyses because of large differences in their day and night y-intercepts, also had the poorest overall level of functioning (lowest y-intercepts) during night flight of all 21 crewmembers studied. Thus, the RG crewmembers who were permitted to take the in-flight nap during night flights were significantly less fatigueable than the NRG not permitted to sleep.

4.5 Physiological Alertness/Sleepiness: Microevent Analysis during Critical Operational Phase

An intensive analysis of specific EEG frequency and EOG changes associated with reduced physiological alertness was conducted on the period from 1 hr. before TOD through landing. This critical phase of operation, including descent and landing, averaged about 90 min. and was analyzed for both the rest and no-rest groups. The entire 90 min. period was scored for the individual occurrence of three specific physiological events: (1) EEG alpha activity (8-12 Hz); (2) EEG theta activity (3-7 Hz); and (3) EOG slow-rolling eye movements (SEMs; > 100 uV amplitude, > 1-sec.

duration). The duration of each microevent occurrence was scored according to three time bins: (1) 5-10 sec., (2) 11-15 sec., and (3) >15 sec.

The physiological microevent data were examined in two ways, and the results of these analyses are presented. First, the raw microevent data were used for an overall descriptive analysis. Second, statistical analyses were conducted in a manner that paralleled the statistical analysis of the PVT lapse data. The specific statistical approach and results are presented.

4.5.1 Raw Data: Descriptive Analysis

The nine subjects in the NRG, each flying four legs, provided a total of thirty-six 90 min. periods. Six of these 90 min. periods were lost because of equipment malfunctions and, therefore, thirty (83%) were available for analysis of microevents in the NRG. The twelve subjects in the RG, each flying four legs, provided a total of forty-eight 90 min. periods. Four of these 90 min. periods were lost because of equipment malfunctions, and the remaining forty-four (92%) were available for analysis of physiological microevents.

The following descriptive analysis of the raw data utilized cumulative totals of the microevent occurrences (a composite score of total alpha, theta, and SEMs microevents). The cumulative total microevents that occurred for all twenty-one crewmembers was 154. The nine NRG crewmembers had a total of 120 microevents (78%), whereas the twelve RG crewmembers had a total of 34 microevents (22%). As expected, most of these microevents, 132 (86%), occurred in the hour before TOD. In the NRG, 98 microevents occurred before TOD, with 22 microevents in the period from TOD through landing. In the RG, all of the 34 microevents occurred before TOD.

There were broad individual differences in the occurrence of physiological microevents. Seven of nine (78%) NRG crewmembers had at least one microevent. Four of these seven (two captains and two FOs) had 9 or more total microevents that together accounted for 84% of the total NRG microevents. Two of these four crewmembers (one captain, one SO on the same trip) accounted for 52% of the total NRG microevents. Six of the twelve (50%) RG crewmembers had at least one microevent occurrence. Two of these six (both SOs) had 9 or more microevents and accounted for 59% of the total RG microevents.

Overall, there were four NRG crewmembers who had more than 11 microevents; an NRG captain had the most occurrences, 42. At the other end of the range, only two NRG crewmembers had as few as 6 microevents. The highest number of microevent occurrences for a crewmember in the RG was 11. Another RG crewmember had 9 microevents and the remaining four RG crewmembers had less than 6 events.

The cumulative total microevents were composed of 87 alpha occurrences (56%), 52 SEM occurrences (34%), and 15 theta occurrences (10%). Most of the microevents were of short duration, 83 (54%) lasting 5-10 sec. Sixty-two (52%) of the NRG microevents were in this time bin, whereas 21 (62%) of the RG microevents fell in this range. Only 23 (15%) of the total microevent occurrences lasted over 15 sec.

The distribution of cumulative total microevents across study flight legs is presented in table 27. It shows that 49% occurred on study leg 4 (a night flight). The NRG had 40% of their microevents on the last study-leg, and the RG had 82% of their occurrences. On study leg 1, there were 21 microevents (18% of NRG) in the NRG and only 1 RG occurrence. Also, most of the microevents, 106 (69%), occurred on night flights (study legs 2 and 4), the NRG having 77 microevents (64%) during the nights and the RG having 29 (85%). However, even here there was tremendous individual variation. The crewmember (NRG) with the most microevents had 37 of his 42 occurrences (88%) during day flights. The crewmember (also NRG) with the next highest total number of microevents had 26 of his 28 (93%) occurrences on the last study flight leg at night. In the RG, the two crewmembers with the highest number of microevents had all of their occurrences on the last study flight leg at night.

Table 27. Raw data-descriptive analysis: cumulative total microevents across study-flight legs

Study flt. leg	RG	NRG	Cumulative totals
1	1	21	22
2	1	29	30
3	4	22	26
4	28	48	76
Cumulative totals	**34**	**120**	**154**

In figure 16, the total cumulative microevents for the NRG (left figure) and RG (right figure) are portrayed in 10 min. time bins across the last 90 min. of flight. As previously indicated, this shows the occurrence of 22 NRG microevents during the last 30 min. of flight (descent and landing phase); all of the RG microevents occurred before TOD.

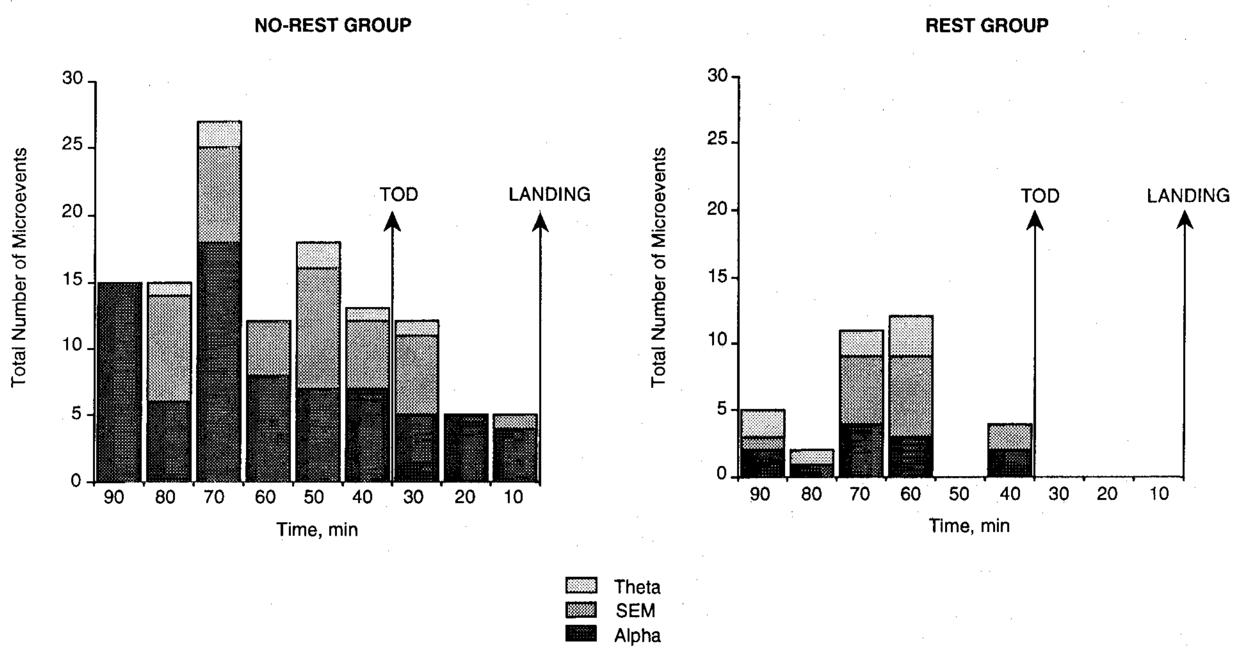


Figure 16. Total cumulative microevents for the NRG and RG portrayed in 10 min. time bins over the last 90 min. of the flight (from about 1 hr. before TOD, through TOD and landing).

4.5.2 Statistical Analysis of Microevent Occurrences

The statistical analysis of the microevents paralleled the approach used to examine the statistical significance of the lapse results from the PVT. As previously stated, the lapses required some consideration before statistical analysis, because there were such large individual differences in lapse frequency among crewmembers. A square root transformation was used on the frequency count of the number of lapses to remove the proportionality between the mean and the variance (ref. 69). This same consideration applies to the statistical analysis of the microevents associated with physiological sleepiness. The occurrence of physiological microevents was quite variable among crewmembers—some individuals had no microevent occurrences, whereas others had many. Therefore, paralleling the analysis of the lapse data, a square-root transformation (square root of x plus the square root of x + 1) was performed to increase the homogeneity of the variance (ref. 69). The purpose of the square root transformation was to reduce the variability of the data set by normalizing the distribution. This normalization reduces the potentially biased representation of specific individuals who had a high number of microevent occurrences.

The ANOVA was then conducted on the transformed data set. Levene's test for equality of variance was examined to determine whether the transformation had normalized the distribution of the data. A significant finding on the Levene's test indicated that a significant difference between the variances remained and that the transformation had not fully normalized the distribution. In this situation, the appropriate nonparametric statistical test was conducted on the raw data (either a Mann-Whitney U rank-sum test for comparison of two factors or the Kruskal-Wallis one-way ANOVA for comparison of more than two factors). The averages (and standard deviations) for the transformed data and the significant findings are presented. Statistical analyses were conducted using the total microevent occurrences (a composite score of total alpha, theta, and SEMs) and also the three separate microevent categories (i.e., alpha, theta, and SEMs).

Overall, the NRG crewmembers averaged significantly more cumulative total microevents, 6.37 (SD = 4.04), that the RG crewmembers, whose average was 2.90 (SD = 2.19) ($F_{1,19} = 6.44$, $p = .02$) (fig. 17). Analysis of alpha, theta, and SEM totals demonstrated a significant difference in the average SEM occurrences between groups. The NRG averaged 3.72 (SD = 2.31) SEM occurrences, and the RG averaged 1.95 (SD = 1.55) ($F_{1,19} = 4.45$, $p = .048$) (fig. 17). A composite total of alpha and theta microevents, without SEM occurrences, also demonstrated significantly more microevents in the NRG than in the RG ($F_{1,19} = 6.38$, $p = .02$).

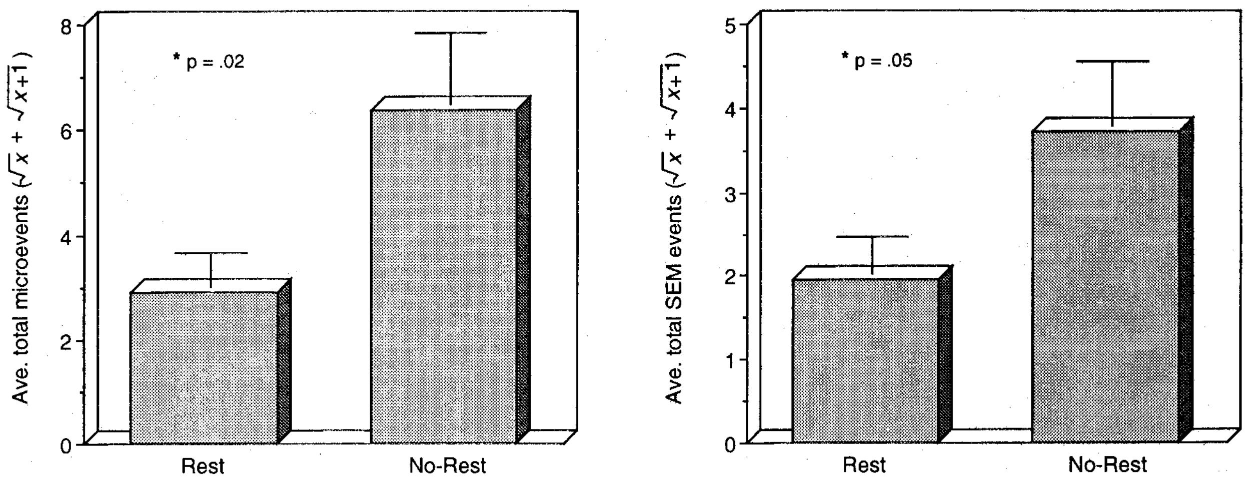
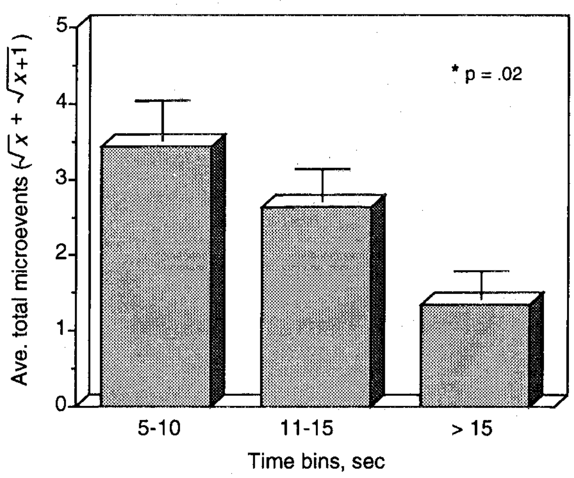


Figure 17. Average number of total microevents occurrences (left figure) and SEM occurrences (right figure) by each crewmember in RG and NRG. Transformed (square root) data are represented in each figure.

There was a significant finding regarding the duration of total microevents, with the most events occurring in the shortest time bin. The average number of total microevents for each of the time bins was 5-10 sec. = 3.43 (SD = 2.42), 11-15 sec. = 2.63 (SD = 1.98), and >15 sec. = 1.35 (0.64) ($F_{1,19} = 4.55$, $p = .015$) (fig. 18). There also was a significant finding for the duration of SEMs, with most events lasting between 5 and 10 sec. The average number of SEMs for each of the time bins was 5-10 seconds = 2.29 (SD = 1.55), 11-15 sec. = 1.70 (SD = 0.99), and >15 sec. = 1.28 (0.90) (Kruskal-Wallis = 7.14, $p = .028$) (fig. 18). A composite total of alpha and theta microevents, without SEM occurrences, also demonstrated the same significant time bin effect ($F_{1,19} = 6.31$, $p = .01$).



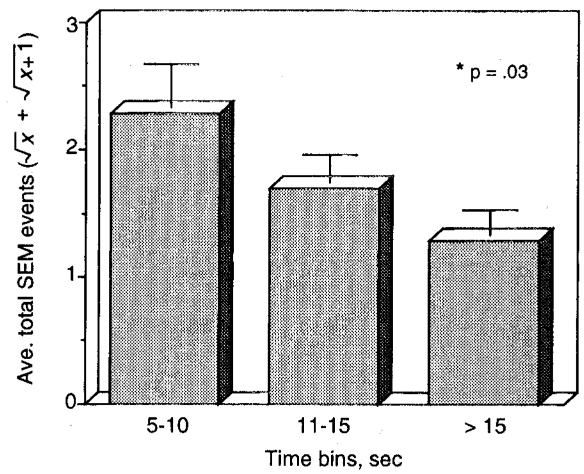


Figure 18. Average number of total microevent occurrences and SEMs for each crewmember of the study group separated by time bins. Transformed (square root) data are represented in each figure.

The period before TOD covered about 60 min., and the period from TOD through descent and landing was about 30 min. Therefore, as expected, the average number of total microevents for the period before TOD, 4.14 (SD = 3.18), was significantly greater than the average number of total microevents from TOD through descent and landing, 1.71 (SD = 1.61) (Mann-Whitney U = 323.5, p = .004) (fig. 19). Also, there were significantly more alpha and SEMs occurrences before TOD. The average number of alpha occurrences before TOD was 2.89 (SD = 2.71); the average from TOD through descent and landing was 1.49 (SD = 1.29) (Mann-Whitney U = 293, p = .026) (fig. 19). The average number of SEM occurrences before TOD was 2.52 (SD = 2.00); the average from TOD through descent and landing was 1.29 (SD = 0.80) (Mann-Whitney U = 292, p =.024) (fig. 19).

Significantly more SEMs occurred on the last study-flight leg. The average number of SEMs by study leg were as follows: leg 1 = 1.07 (SD = 0.31), leg 2 = 1.70 (SD = 1.22), leg 3 = 1.21 (SD = 0.98), and leg 4 = 2.02 (SD = 1.64) (Kruskal-Wallis = 9.76, p = .02) (fig. 20).

Two significant findings emerged for microevent occurrences on day versus night flights. The average number of total microevents on each night flight, 2.52 (SD = 1.72), was significantly greater than the average on each day flight, 1.65 (SD = 1.71) (Mann-Whitney U = 694, p = .035) (fig. 21). Also, the average number of SEM occurrences was significantly greater on night flights, 1.86 (SD = 1.09), than on day flights, 1.14 (SD = 0.51) (Mann-Whitney U = 652, p = .002) (fig. 21). A composite total of alpha and theta microevents, without SEM occurrences, also demonstrated significantly more microevents on the night flights than on the day flights (Mann-Whitney U = 676, p = .02).

4.5.3 Sleep Latency Results

As indicated earlier, the speed of falling asleep (sleep latency) is an accepted laboratory measure of physiological sleepiness, increased sleepiness being associated with shorter sleep latencies (i.e., falling asleep quickly). The laboratory standard for a level of excessive physiological sleepiness is a sleep latency of 5 min. or less, sometimes referred to as the "twilight zone" (refs. 21, 27). In this study, the RG average was 5.6 min. to fall asleep. (For the RG (24) subjects this average was 4.1 min.) This indicates that overall, this group of volunteers fell asleep quickly and close to the laboratory range that indicates excessive physiological sleepiness.

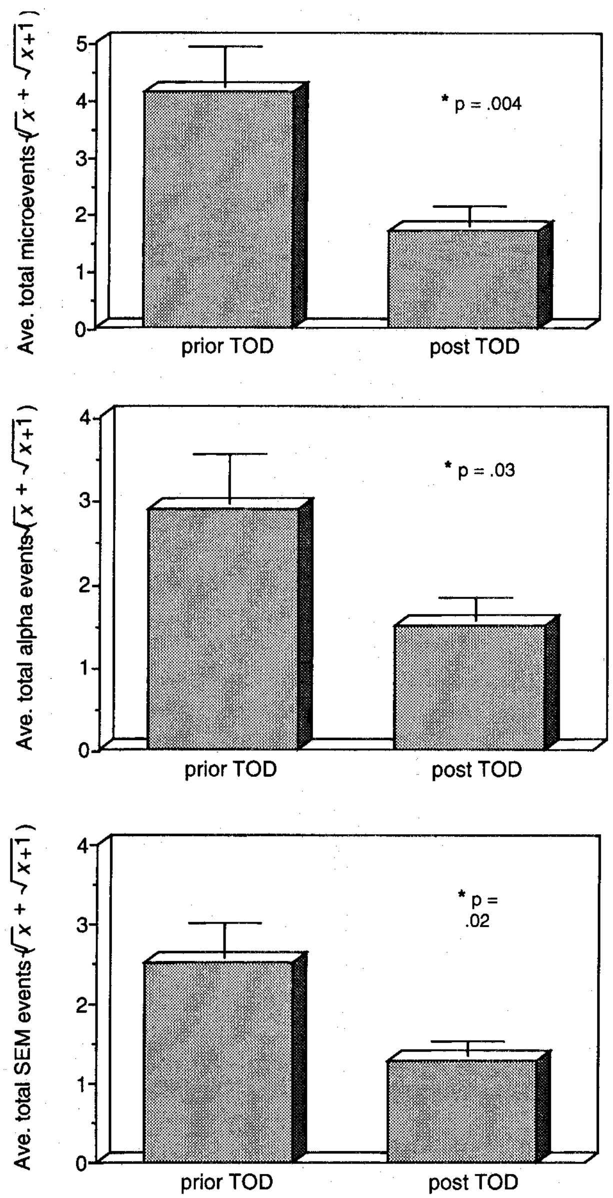


Figure 19. Average number of total microevent occurrences, alpha, and SEM for each crewmember of the study group separated by prior-to TOD and following (TOD) for each study leg. Transformed (square root) data are represented in each figure.

4.5.4 Subjective Alertness Ratings

Crews rated their alertness every hour throughout each flight leg. Following each flight they retrospectively rated their overall alertness during the flight. Both types of subjective ratings were collected on 10-cm analogue scales rated from most drowsy to most alert. Because of missing data, in-flight alertness ratings were subdivided into the average of two ratings immediately before and after the rest/control period. Thus, the three-way ANOVA on self-reports of in-flight subjective alertness was structured to include both conditions (rest vs. no-rest), four flight legs, and two phases within each flight leg (pre-control vs. post-control).

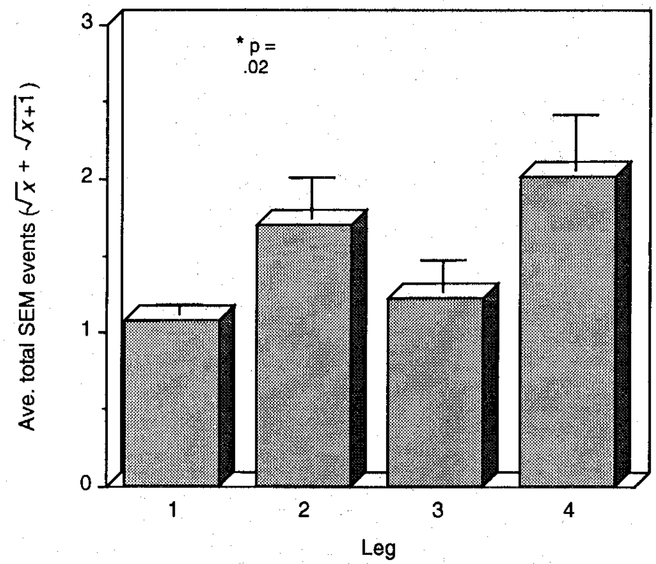


Figure 20. Average number of total occurrences of SEM for each crewmember of the study group over each study leg. Transformed (square root) data are represented.

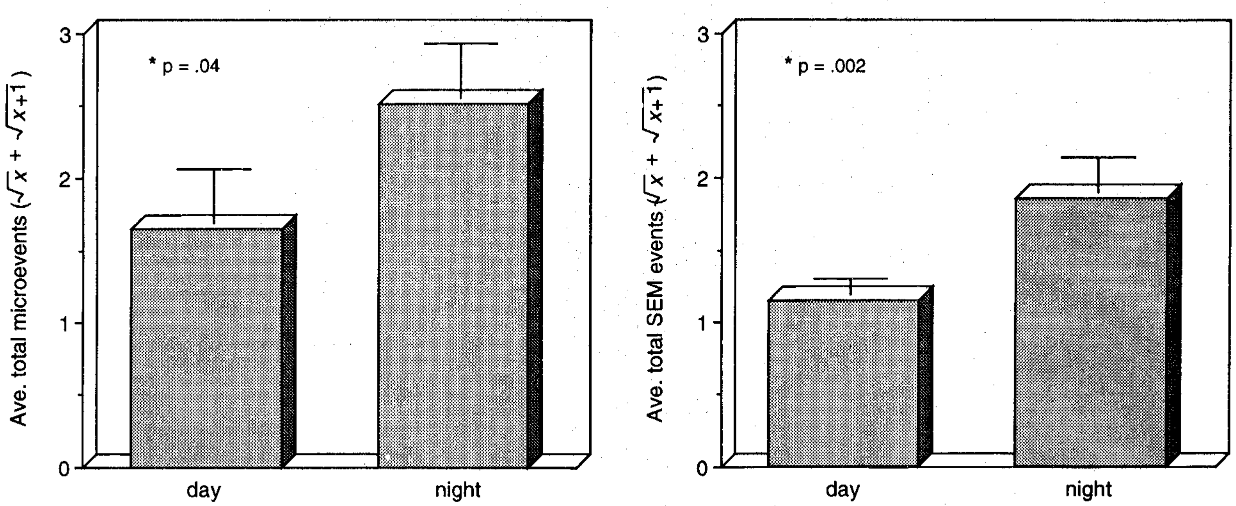


Figure 21. Average number of total microevent occurrences and SEMs for each crewmember of the study group per day and night-flight legs. Transformed (square root) data are represented in each figure.

The analysis of alertness ratings yielded a main effect for flight leg ($F_{3,48} = 19.2$, $p < .0005$). Not surprisingly, subjective alertness was lower on night flights (legs 2 and 4) than on day flights (legs 1 and 3). There was also a main effect for phase of flight ($F_{1,16} = 28.8$, $p < .0005$), resulting from post-rest alertness ratings being lower than pre-rest ratings. However, this varied with flight leg, yielding a significant leg-by-phase interaction ($F_{3,48} = 12.8$, $p < .0005$); subjective alertness decreased from the pre-rest to the post-rest phases of flight for flight legs 2, 3, and 4, but not on flight leg 1 (fig. 22).

The nap did not appear to affect subjective alertness ratings, although there was a trend for the RG to average higher alertness ratings overall than did the NRG (condition main effect, $F_{1,16} = 3.9$, $p = .063$). The nap did not interact with flight leg or phase of flight, but there was a significant interaction among condition, flight leg, and phase of flight ($F_{3,48} = 3.1$, $p = .033$). On flight leg 1, the nap resulted in increased alertness ratings in the RG, whereas all other flight legs for both groups resulted in comparable decreases in subjective alertness across the rest/control period. Consequently, there was no systematic evidence that the nap altered the decrease in subjective alertness experienced as time passed on a flight leg. An analysis of postflight reports of alertness also showed no differential effects of the nap on alertness, although like their ratings of in-flight alertness, crews reported themselves significantly less alert on night-flight legs (2 and 4) than on day-flight legs (1 and 3) ($F_{3,33} = 5.8$, $p < .002$).

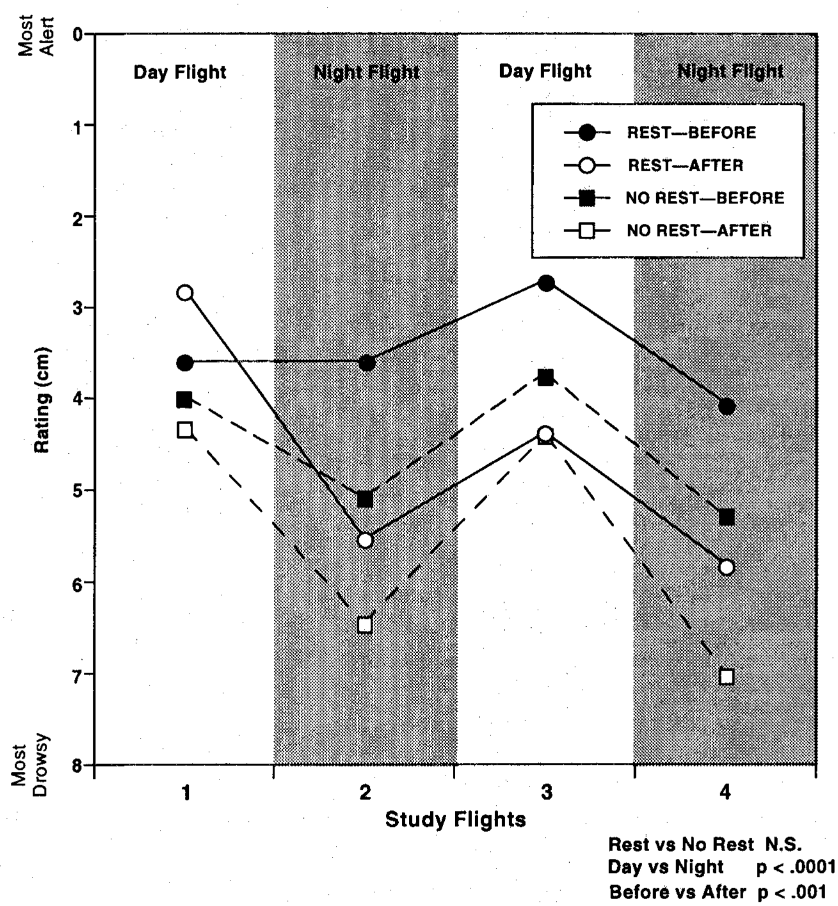


Figure 22. Mean subjective alertness ratings for subjects in RG and NRG conditions for each of the four flight legs.

Thus, in-flight naps were associated with improved performance and physiological measures of alertness. However, they were not associated with decreased subjective fatigue or improved subjective alertness. This failure to find naps affecting subjective activation may be due to the way in which data were averaged because of missing data, but the finding is also consistent with results from laboratory studies of naps taken during periods of sustained wakefulness (ref. 14). In those studies, naps clearly enhanced performance and physiological activation, but did not change subjective activation. Although the reasons for this finding in the current study are unclear, it is further evidence that subjective reports from flight crews do not always reflect accurately the level of physiological sleepiness that may be present (refs. 39, 40). As expected, crews did report decreased alertness on night flights, and lower alertness as time progressed within each flight, and these two effects were robust. However, against the backdrop of such temporal effects, in-flight naps did not result in subjective changes comparable to those recorded in the objective indices of alertness.

4.6 Layover Sleep: Results from Wrist Activity Monitor and Sleep Log

Each crewmember wore a wrist activity monitor (WAM) continuously for 1 to 5 days before the duty cycle, while on duty and throughout each layover, and for 1 to 3 days after the duty cycle. The actigraph provided objective documentation of crewmembers' self-reported sleep episode. Used in this way, the actigraph helped validate the sleep log and allowed "correction" of the self reported layover sleep amounts. Thus, if a crewmember reported sleeping at a time when the actigraph was showing active motility identical to wake ambulation, then the sleep report was discounted. These rather frequent errors usually involved a crewmember apparently misrecording or misperceiving the time of a sleep episode (e.g., the actigraph indicated that a layover sleep began at 0645 GMT, but the crewmember logged it at 0845 GMT)—these kinds of errors are common in retrospective reports. Thus, the actigraph permitted an objective check on times when crewmembers reported sleeping, but it could not guarantee that sleep was actually present during a period of low motility (especially a short-duration period) and it could not provide data on the stage of sleep at any given time. For these latter goals, polysomnography is necessary.

Following detailed correction of the sleep log using actigraphic information, the timing of layover sleep and the cumulative sleep loss across duty cycle days were calculated for each crewmember. The analysis of the data focused on two questions, the first of which was concerned with the extent to which the cockpit naps alter the sleep debt and layover sleep patterns, and the second with individual differences in sleep debt.

4.6.1 Cumulative Sleep Loss and Cockpit Rest

Figure 23 displays actigraphic rest/activity patterns from a captain during two phases of the study. The top panel shows his motility during the 73 hr. period at home, immediately before beginning a duty cycle; the bottom panel shows his motility pattern for the second 73 hr. period of duty (during days 4, 5, and 6). The pattern and timing of sleep (black horizontal bars in the figure) changed during duty, but there is also a clear decrease of 6% (4.38 hr.) in the proportion of time occupied by sleep. The patterns for trip leg 2 (duty days 1, 2, and 3) and trip leg 4 (duty days 7, 8, and 9) are similar to the bottom panel of figure 23 and illustrate clearly that this NRG captain developed a cumulative sleep debt during the study.

The development of a cumulative sleep loss portrayed in figure 23 was displayed by most of the crewmembers in the study. Similar to the results of an earlier study examining long-haul crews flying polar routes (ref. 70), not all crewmembers developed a pattern of cumulative sleep loss during a duty cycle. Sasaki and his colleagues reported that 10 of 12 long-haul crewmembers (83%) suffered a cumulative sleep debt during operations, that is a debt of at least 4 hr. by the ninth day, with the worst case reaching 25 hr. of lost sleep by the ninth day. The results from the current study are very similar. Eighteen of a total of 21 crewmembers (86%) had a cumulative sleep debt of at least 4 hr. by the ninth day, with the worst case reaching 22 hr. When combined with the earlier study, it suggests that 85% (28/33) of long-haul crewmembers develop a cumulative sleep debt after repeated days of transmeridian duty.

Sasaki et al. (ref. 70) did report that two long-haul crewmembers (17%) actually gained at least 4 hr. of sleep by the ninth day of a duty cycle, but this occurred in only one (5%) crewmember in the present study. Two other crewmembers had neither a sleep debt nor a gain. Interestingly, the three crewmembers without a cumulative sleep debt in the study also reported relatively short periods of sleep daily while off duty at home. In fact, the crewmember who gained sleep by day 9 of the duty cycle (+8.5 hr.) reported the least sleep at home off duty (6.5 hr./day). The pattern for the other two crewmembers (no sleep debt) was similar: the first had +0.6 hr. cumulative gain by day 9, and reported 7.0 hr./day of sleep at home; the second had -1.0 hr. cumulative loss by day 9, and reported 7.3 hr./day of sleep at home. Nevertheless, the vast majority of crewmembers developed a sleep debt as the study progressed.

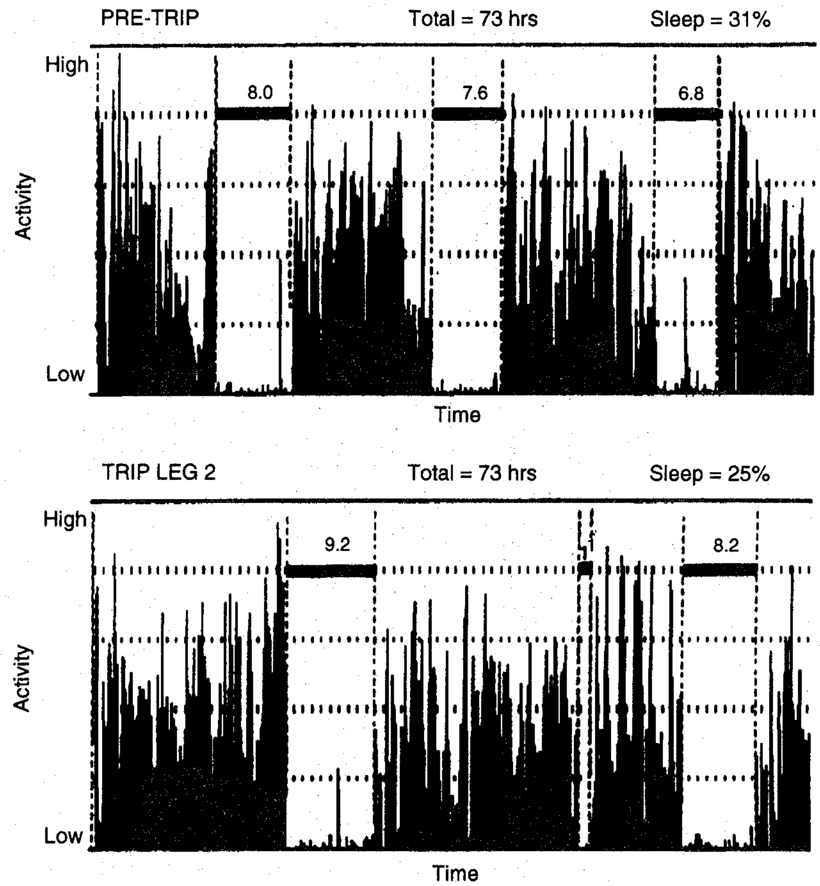


Figure 23. Wrist actigraphic recordings from a NRG captain during two phases of the study. Top panel shows rest/activity pattern during the 73 hr. period at home, immediately before beginning a duty cycle. Bottom panel shows rest/activity pattern for the second 73 hr. period of duty (during days 4, 5, and 6). Periods of sleep are represented by horizontal bars. The proportion of sleep time reflected in each record is shown the upper right-hand corner of each panel.

To determine whether the cockpit rest altered the cumulative sleep debt of crewmembers, analyses examined the cumulative sleep debt of the RG subjects both with and without the cockpit naps included, as well as sleep debt of the NRG subjects. Figure 24 displays the average sleep debt for these three conditions and the first-order linear regressions fitted to the data from each condition. There was no statistically significant difference between the NRG and the RG in cumulative sleep loss functions. In the RG, when the cockpit nap is not included, there was a trend toward a greater sleep debt. This appeared to be largely a result of the RG sleeping somewhat less than the NRG during the layovers in Honolulu and Los Angeles. There is no evidence that this was related to naps in-flight, for it did not occur at other layovers.

It is important to highlight that the cockpit nap did not have a significant effect on the cumulative sleep loss function of the RG. This suggests that whatever benefits the nap had for performance and alertness, those benefits were not created by a diminution of the cumulative sleep loss experienced by crews. By the ninth day of duty, with or without planned cockpit rest, crewmembers averaged approximately 1 full night of lost sleep. Rather than attenuating this more chronic source of fatigue in long-haul operations, it appears that the cockpit naps functioned as an acute relief for fatigue, promoting alertness but not affording enough sleep to circumvent accumulated loss over many days.

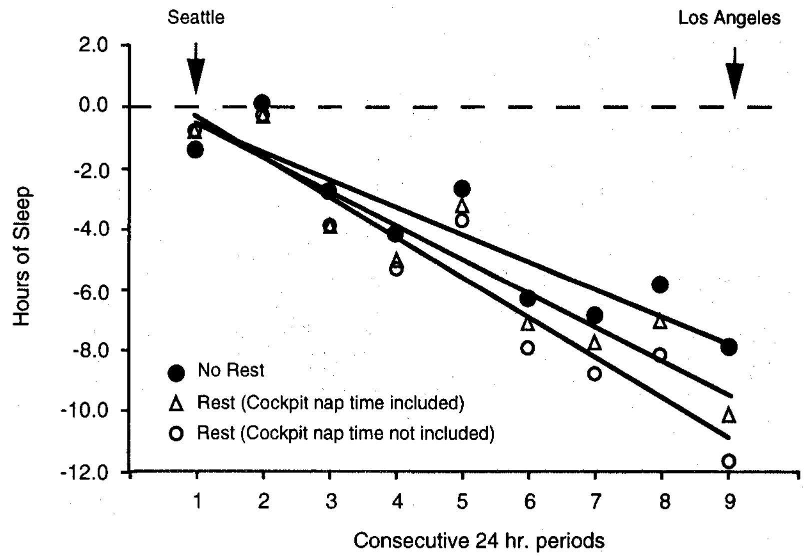


Figure 24. Mean hours of cumulative sleep loss (by combined actigraph and sleep log data) during nine consecutive 24 hr. periods of the study duty cycle (through the LAX layover). Data are shown for the NRG and RG with and without the cockpit rest periods included. Linear regression functions fitted to each group's data are also displayed.

4.6.2 Layover Sleep Episodes

Two findings, the accumulated sleep debt developed by most crewmembers and the lack of an associated effect from the cockpit naps, prompted an evaluation of the timing of layover sleep episodes. This was examined for the seven approximately 24 hr. layovers of the study duty cycle. The assumption commonly made is that a 24 hr. layover should provide adequate opportunity for rest and sleep. Yet as demonstrated above, the vast majority of crewmembers developed a sleep debt. Figure 25 displays the average percent of time spent asleep (by combined actigraph and sleep log) at each layover for both RG and NRG crewmembers combined. On four of the seven layovers, crews averaged about 40% sleep time, which is comparable to 9.6 of 24 hr. This is not an inconsequential amount of sleep, and it suggests that crews were endeavoring to obtain reasonable amounts of sleep on layovers.

Both RG and NRG crewmembers obtained about 5% less sleep at two layovers (Osaka and Narita2), for reasons that are not yet clear. Osaka was the longest duration layover (29.4 hr.—see table 1), which may have resulted in proportionately less sleep being obtained. Similarly, both groups obtained the highest proportion of sleep on the Los Angeles layover, spending an average of 45% (10.8 of 24 hr.) sleeping. There are three factors that may have contributed to this increase in layover sleep at LAX: (1) the Los Angeles layover was on home time for all but a few of the crewmembers; (2) it occurred near the end of the duty cycle, when a sleep debt had already developed for most crewmembers; and (3) it occurred immediately before the flight with the longest duty time (LAX to SEL). Consequently, when reaching Los Angeles, crews were tired, they were sleeping at times that were consistent with their home circadian cycle, and they were aware that the next trip leg would involve the longest duty duration of the trip.

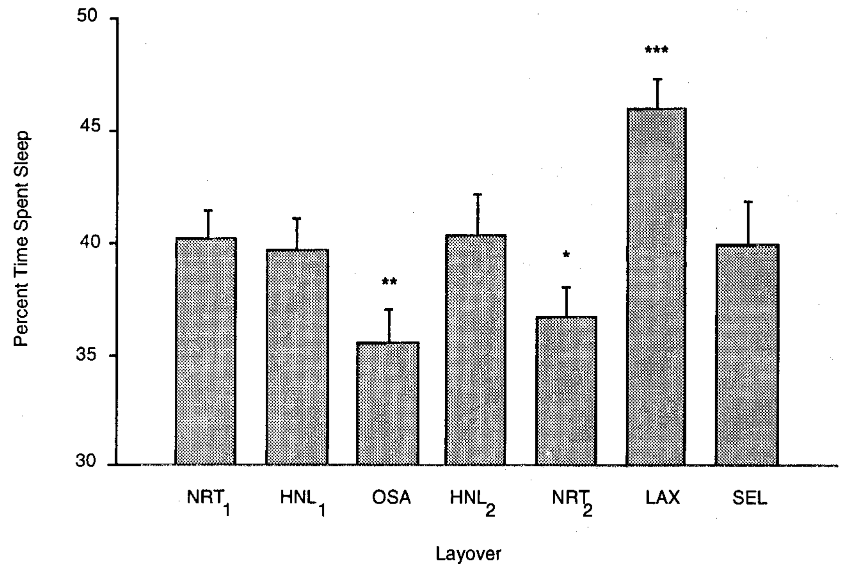


*Figure 25. Mean (S.E.) proportion of time spent asleep (by combined actigraph and sleep log) at each of seven duty cycle layovers for all 21 crewmembers (RG and NRG). Asterisks highlight layovers with sleep percentages significantly below or above the 40% levels of NRT1 and HNL1 (*p <.05, **p < .01, ***p < .001).*

The relatively high proportion of sleep time on layovers appears inconsistent with the sleep debt that was accumulated by most crewmembers. However, the inconsistency disappears when the full duty work-rest cycle is considered. Figure 26 displays the average proportion of time spent asleep 1) when crews were off duty, at home; (2) when they were on layover; and (3) when layover was combined with subsequent duty time. The 24 hr. day at home can be subdivided into a work:rest ratio of 1:2. This represents the 8 hr. of the day that are typically devoted to work (even when not strictly working) and the remaining 16 hr. devoted to rest (to include exercise, eating, social activity, sleep). At home, crews spend about 33% of the 24 hr. day sleeping, whereas on the study layovers they slept about 40% of the time. This suggests that they used the layover time for more sleep than they typically obtained at home, off duty. But when layover rest time (about 24 hr.) is added to subsequent duty work time (about 12 hr.), the result is a 36 hr. duty period. This represents roughly a 1:2 work:rest ratio (12 duty hr.:24 layover hr.). For crews in the current study the average duty period ratio ranged from 1:1.30 (SEL) to 1:2.79 (HNL2). Across 36 hr. duty periods, the average proportion of time crews spent asleep was 28%, which is 5% below what they typically obtained at home on a 24 hr. day. Consequently, when unencumbered by work during a 24 hr. layover period, crews slept proportionally more than when at home, but when the full 36 hr. work-rest cycle or duty period is considered, the proportion of sleep is significantly less than that obtained on a 24 hr. day at home. Hence, the increased amount of layover sleep obtained by crews is offset by their long duty period, resulting in a net loss of sleep for most crewmembers. It is unclear whether these results are unique to 24 hr. layovers, or whether cumulative sleep loss accrues for most crewmembers flying any long-haul trip schedule that involves layovers.

Finally, an analysis of the layover sleep episodes also revealed that approximately 40% of the sleep time was generally not obtained in one sleep period during the layover. Out of a total of 135 layovers, 77% involved two or more sleep episodes. Figure 27 displays the histogram of sleep episodes on layover. Most layovers involved two sleep episodes. There was a striking relationship between the duration of each sleep episode within a layover, as depicted in figure 28. Layover sleep durations were negatively correlated ($r = -.82$, $p < .0001$), such that when the first sleep episode in a layover was 6-11 hr. in duration, the second sleep episode either did not occur, or if it did occur, it was under 4 hr. in duration. Conversely, when the first sleep episode was under 6 hr., a second sleep episode virtually always occurred and was between 4 and 11 hr. in duration. The fact that layovers tended to include two sleep episodes appeared to reflect a compromise between the influence of local layover time (e.g., food availability, quiet) and the influence of preferred circadian phases for sleep (ref. 3). A later report will review the timing of these layover sleeps and their relationship to crewmembers' home time.

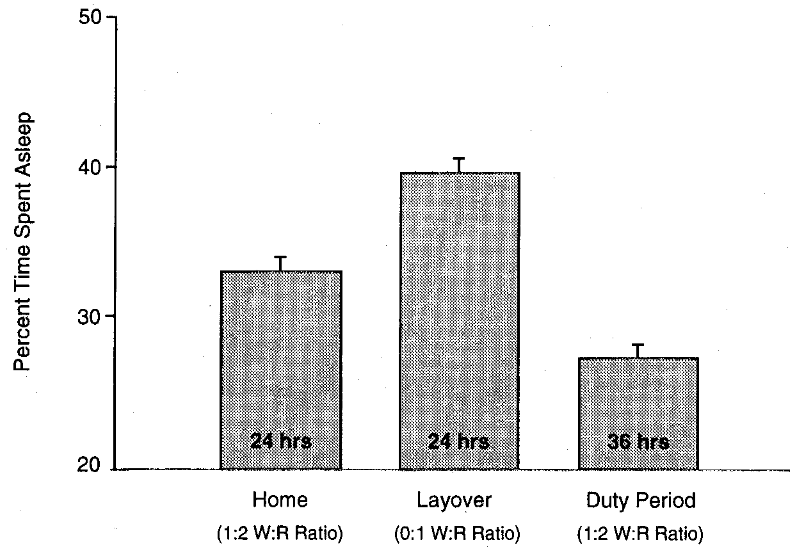


Figure 26. Mean (S.E.) proportion of time spent asleep when crews were off duty (home), when they were on layover (layover), and when layover was combined with duty time (duty period). W:R refers to ratio of time typically devoted to work (W) to that devoted to rest (R) within each time-frame. The duration of each time-frame is shown within the bottom of each histogram.

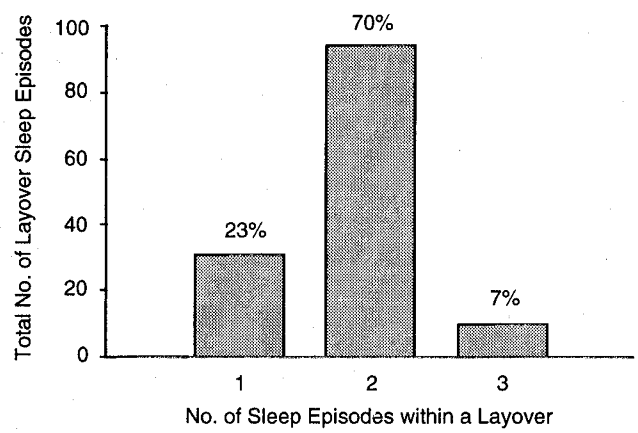


Figure 27. Total number of sleep episodes (by combined actigraph and sleep log) on layovers as a function of frequency of sleep episodes within each layover for all 21 study crewmembers (RG and NRG). Percentages reflect the proportion of each of the grand total of layover sleep episodes.

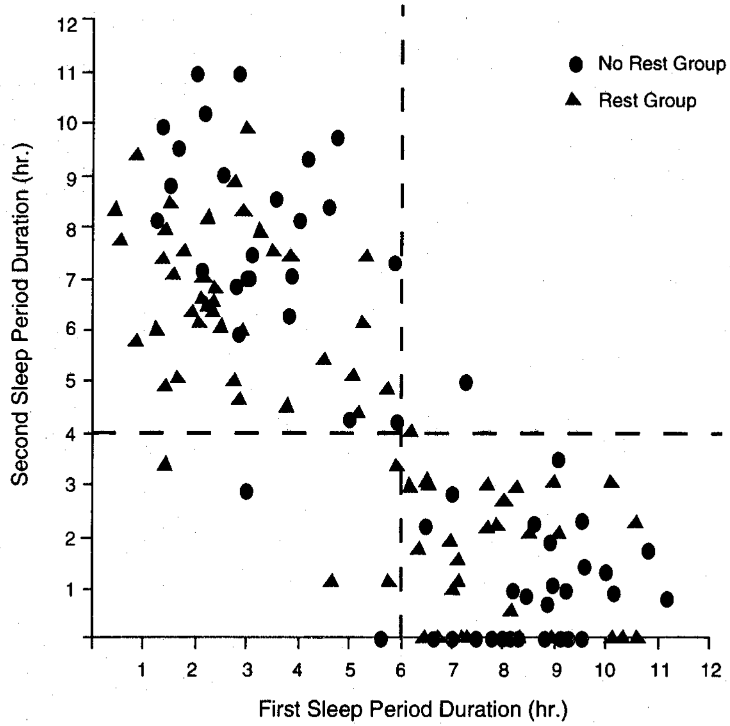


Figure 28. Duration of second sleep episodes on layover as a function of the duration of first sleep episodes for RG and NRG.

5.0 DISCUSSION

5.1 Study Limitations

This study involved only one trip pattern on a commercial airline. The study trip pattern was chosen according to predetermined criteria, but clearly the variety of trip schedules currently and potentially available is tremendous. Also, it is difficult to assess how the specific airline cultures may have affected the study outcomes. The study was conducted on transpacific flights to utilize the opportunity of scheduling the planned rest periods during cruise over water. Therefore, the low-workload portion of flight identified in this study occurred over water. The intense physiological and performance data collection occurred during a specific and restricted middle segment (four consecutive flight legs) of the trip schedule. Therefore, the initial home-to-flight-schedule transition is quantified only with logbook and actigraph data. Also, the final trip legs, which may represent the highest levels of accumulated fatigue, were not studied except for logbook and actigraph data. This study involved B-747 aircraft flown by three-person crews. Questions have already been raised regarding the applicability of this study to the two-person cockpit. There were two NASA researchers on the flight deck during the in-flight data collection periods. Although they were instructed to minimize their interactions with the crew and to make their presence on the flight deck as unobtrusive as possible, there is no question that having two extra persons in the cockpit may have potentially altered the regular flow of cockpit conversation and interaction. It is important to remain cognizant of these limitations when generalizing these results. As always, it is not appropriate to generalize the study results to scenarios that extend beyond the scope of the specific scientific issues addressed here.